Hyperspectral tissue pre-screening and segmentation for enhanced raman-based oral cancer diagnosis

Jarman, Anneliese

Awarding institution:
King's College London

Hyperspectral Tissue Pre-Screening and Segmentation for Enhanced Raman-Based Oral Cancer Diagnosis

Anneliese Laura Jarman

October 18, 2021

A thesis submitted for the degree of

Doctor of Philosophy

Faculty of Dentistry, Oral & Craniofacial Sciences, King's College London

Acknowledgements

This started with flash of incredible serendipity: a discussion between a child and two adults about the apparent size of a pickle in a vinegar jar in a fish and chip shop.

This thesis is dedicated to The Pickle.

I must start by thanking both of my supervisors, Prof. Richard Cook and Dr. Susan Cox, both for their shared expertise spanning many fields and for their practical support in navigating so many unforeseen challenges. I am a better scientist for working with you both.

The generous contributions of the tissue samples, their preparation and the key gold standard data for my cross-modal comparison by Dr. Selvam Thavaraj and Miss Rhonda Henley-Smith, along with Prof. Richard Cook, were invaluable. Witnessing their knowledge of such an expansive field, first-hand, fed my interest in digital pathology throughout this project.

The opportunity to also work with Dr. Stephanie Wright and her group on a project straddling health and environmental themes was so rewarding and I'm particularly grateful for the continuation of our collaboration after her departure from KCL to Imperial College.

Thank you to Renishaw for their contributions to this project, including the continued advice and support of Dr. David Reece and Dr. Martin Isabelle.

Aside from these vital contributors, my project has been bolstered by so many magnanimous colleagues. The whole of the Cox group have contributed their ideas in group meetings over the years, with particularly practical help afforded by Dr. Richard Marsh and Dr. Liisa Hirvonen and proof of principle samples prepared by Alex Gorey and Tom Phillips.

Thank you to my friends and colleagues on Floor 17 over the years for their snippets of advice and knowledge and for brightening the long days in the microscopy cave. Floor 17 would not run without Pete Pilecki; thank you for tolerating my first foray into microscopy in your lab so many years ago and chasing down my orders ever since. Thank you to Dr. Victoria Boyes for saving my life.

I would never have made it here, in so many ways, without my Mum: My greatest supporter, from my very first tentative steps towards a career in science to this day. Thank you.

I'd like to end by thanking Dylan for picking me up and dusting me off for years and, finally, for carrying me over the finish line.

Abstract

Oro-pharyngeal cancer incidences have been increasing in recent years and late detection means that the prognosis is often poor. In spite of under-availability of trained histopathologists across the United Kingdom, there has been very little clinical translation of automated, *in vivo* diagnostic devices, in spite of their proven sensitivity. This is potentially due to their requiring widespread change to the established, gold standard diagnostic workflow. A device which could be used to support histopathology in the detection of cancer from thin tissue biopsy sections may be more easily adopted. Raman spectroscopy has been identified as a highly specific diagnostic tool but it is an extremely time consuming technique which has prevented simple clinical application and translation. In order to make Raman a realistic diagnostic aid to histopathology, a rapid pre-processing technique is required to identify regions of interest to accelerate and streamline its application. The feasibility of hyperspectral imaging (HSI) for this purpose is investigated in this project. Three systems were built consecutively; a preliminary, an improved and a more economical system, respectively. The preliminary system, utilising a Halogen source, was used for proof of principle with data collected from fluorescent dyes, blood, stained oral tissue, saliva and unstained oral tissue. Using this range of absorbent and non-absorbent, biological and non-biological samples, spectral accuracy and variations with concentration were demonstrated. Additionally principal component analysis (PCA) data denoising and k-means clustering of these hypercubes was shown to be successful, to varying degrees, even in colourless tissue samples, chiefly being sensitive to fibrous tissue and changes in cell density. The improved system garnered data with an increased SNR by utilising a powerful white light laser and with the inclusion of a comprehensive background removal protocol. Segmentation of malignant tissue with PCA denoising and k-means clustering showed similarities with histopathologist cancer selection on corresponding H&E stained tissue sections. Several alterations to the system and software were made in order to facilitate coregistration, and quantitative comparison, between the three modes; histopathology, hyperspectral imaging and Raman spectroscopy. This final setup was built to be cost-effective for clinical appeal, with a pulsed Xenon source and fibres replacing a white light laser, a beam expander and a number of lenses. This system covered a wavelength range of 521 - 899 nm (6 nm spectral resolution) and performed with an AUROC (area under receiver operator curve) score of 0.70 on the pixel-wise segmentation of tongue squamous cell carcinoma in frozen, unstained tissue sections compared to consensus histopathology-based diagnosis, with a 2 µm spatial resolution. It performed similarly in comparison with Raman spectroscopy, garnering an AUROC score of 0.69, suggesting that their cancer segmentation is reasonably compatible and they could be used to good effect in a combination system. The comparison

and evaluation process was vulnerable to tissue sectioning artefacts, which are unavoidable, and random fluctuations in illumination intensity and suggestions to limit the impact of these weaknesses in future work are suggested. The system and quantitative evaluation process requires very little user input and could facilitate easy development of this system, and others. This automation, and the fact that it can be used on thin, unstained tissue sections means that this system has the potential to dovetail neatly with the current histopathology-based protocol. With the suggested improvements, this hybrid HSI-Raman system could provide a second opinion, and potentially evolve to share the load and expedite the oral cancer diagnostic process and improve survival rates.

Contents

List of Figures

List of Tables

1 Introduction

Oral cancer is a worsening problem, both in the United Kingdom and globally, but its poor survival rate is improved by nearly 100% if it is detected early, before it has metastasized.[1] Whilst reliance on self-referral by an under-informed public is the primary bottleneck in the diagnostic procedure, a nationwide shortage of histopathologists is now a significant contributing factor. Compounding this issue, there remain significant time delays in the post-biopsy timeline. To ameliorate the oral cancer prognosis, the scientific community has designed several point-of-care devices for oral cancer screening in routine dental appointments. Unfortunately none of these devices have been widely accepted into NHS clinics as they require substantial deviation from the established gold standard biopsy and histopathological diagnostic procedure, and from routine dental check-ups, in addition to requiring clinicians to receive additional training. An automated, quantitative technique which works on unstained, biopsied tissue sections may, therefore, have a higher uptake due to its suitability for use by an unskilled operator and with the resulting in a reduction in negative checking time and the added benefit of increased reliability. To that end, a combination hyperspectral scattering microscopy and Raman spectroscopy system is proposed, exploiting the speed and high data dimensionality of one technique and the diagnostic specificity of the other.

1.1 Oral cancer

In the United Kingdom the annual incidence data indicates that over 8700 cases of oral or oropharyngeal cancer were diagnosed in 2020, with almost a third of those patients losing their life to it.[2] This places it in the top 15 most prevalent and most deadly cancers in the UK, and in the top 10, globally.[3] In spite of ever-growing knowledge and understanding of oral cancer, its risk factors and its presentations, its prevalence in England has increased by 58% in the last decade and 97% in the last 20 years, partly shown in Figure 1 (reproduced from [2]).[2]

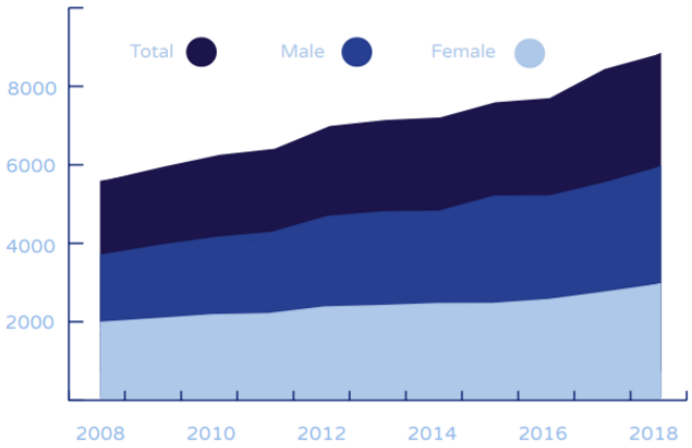

Figure 1: Mouth cancer incidence in the UK, year on year from 2009 - 2019, inclusive.

The ten year survival rate is strongly dependent on the cancer site, lesion size and aggression (characterised by growth and cancer grade), its invasion (specified by cancer stage) and rapidity of diagnosis. Early diagnosis has the single largest impact on outcome, far beyond any technical development in surgical, radio- or chemotherapeutic technologies. This, alone, can improve chances of survival from 50% to 90%, as shown in Figure 2 (reproduced from [2]).[4, 5]

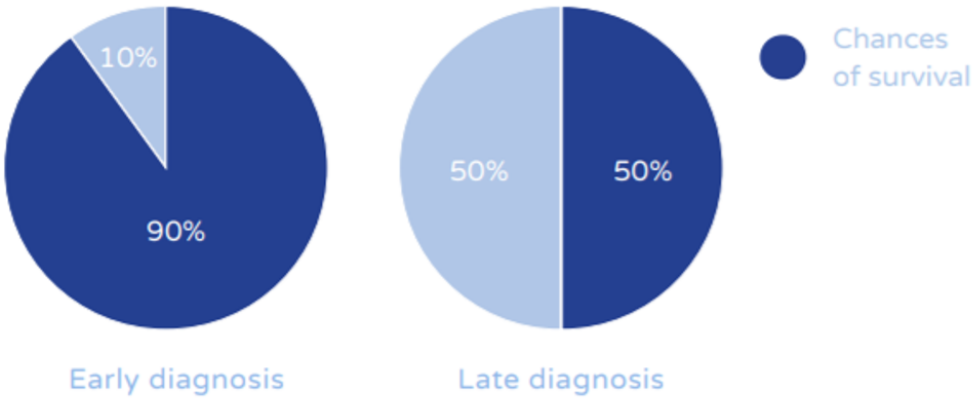

Figure 2: Chances of survival based on early and late diagnosis.

The diagnostic gold standard on discovery of a suspicious oral lesion is a biopsy; the surgical excision of one or more representative samples of the lesion's tissue, which is fixed and sent to a histopathology lab, where it is traditionally wax embedded (after decalcification if the sample contains mineral), thinly sectioned, stained and examined by a clinical pathologist.[6] Unfortunately, self-referral is the most common first step in this diagnostic pathway, followed by dentists identifying suspicious lesions during regular dental appointments. This reliance on self-referral requires that national understanding of oral cancer is of a high level but unfortunately, while it is improving, it

is currently very poor with 83% of British adults not knowing the signs and symptoms of oral cancer.[2] Due to ignorance of the disease, fear of oral healthcare and of cancer, reliance on self-referral is arguably the greatest contribution to the overall diagnostic bottleneck.[7] Nonetheless, any potential reduction in overall lead time to establishing diagnosis and a commitment to treat (or diversion of the patient away from cancer services in the face of a benign diagnosis), is essential to improve outcomes further.

Correctly diagnosing oral malignancies is vital to ensure timely, suitable treatment and appropriate aftercare. However, clinical visual inspection alone is woefully inadequate as a single simple screening process. Oral mucosal lesions are notoriously fickle, being difficult to distinguish from one another and with some benign-appearing malignant lesions and vice versa.[8] This mandates a significant reliance on specialist oral pathology services to characterise lesions accurately. Further complications arise with the increasing recognition of potentially pre-malignant lesions, which may proceed to a cancer at some stage; the hierarchy of risk increasing from leukoplakias (white patches with no other aetiopathological[1] explanation), through mixed erythroleukoplakias to erythroplakias (red patches with no other aetiopathological explanation) which are the least stable of all possible lesions.[9] In addition, more recently recognised lesions such as proliferative verrucous leukoplakias and the genetic pre-disposition syndromes such as Fanconi's aplastic anaemia and dyskeratosis congenita (in which 1/3 of children diagnosed will have an oral cancer by their 30th year), further complicate the surveillance and screening exercises. A visual diagnosis of either of these conditions in clinic is nearly impossible and so depends only on other inflammatory lesions such as endogenous inflammation (lichen planus and lupus erythematosus) being discriminated from the traumatic lesions such as morsicatio (chronic chewing of the tongue, lip or cheek), by histopathology. The umbrella diagnostic terms and key diagnostic issues in oral health are the discrimination of benign tumours and lesions (such as pyogenic granulomatous epulides, giant cell disorders and torus palatinus), oral ulcerative erosive and immunobullous disorder groups, direct traumas such as cheek or tongue bite, infections such as bacterial, viral or mycological stomatitis, periodontal diseases, haematological impacts (such as haematinic deficiencies driving glossitis and leukaemias/lymphomas driving gingival lesions), leukoplakia and erythroplakia and auto-immune disorders (such as lupus and oral lichen planus).

To this end, awareness of patterns in malignant transformation potential are also important and represent another part of the complex diagnostic picture which informs pathologists in diagnostic and monitoring decisions. Predominantly, even if the precursor lesion appears benign, long term inflammation is more likely to undergo malignant conversion in sites of phenotype change. Examples of these sensitive locations include

[1] Aetiopathology is the study of the cause of a certain pathology.

the corner of the lips where inner mucosal and external skin types abut, and lateral to ventral tongue where the specialised papillated dorsal mucosa alters to become the thin smooth under-surface.

Squamous Cell Carcinoma (SCC) is defined as an uncontrolled growth of the squamous cells of the epidermis due to mutations in their DNA and it accounts for about 94% of malignancies of the oral cavity, being the most common metastasising cancer from this region.[9, 10] As a result it is very dangerous although, as noted above, its high morbidity and mortality rate is largely due to late detection, degrading the prognosis more powerfully than nearly all the other lesion characterisations. The time between the patient first experiencing symptoms to their visiting a health care professional has a mean time of 104.7 days with a notably large range of 0-730 days.[11]

Over half of oral and oropharyngeal cancers are found on the tongue or tonsils, both of which would typically be squamous cell carcinomas. Whilst the decreasing popularity of smoking has led to decreases in oral cancer over the last few decades, the incidence of tongue squamous cell carcinoma has increased.[12] This increase, especially in cancers on the distal third of the tongue which fall into the category of oropharyngeal cancer, is linked to increased sexual transmission of the human papilloma virus (HPV) which is being combatted in part by the nationwide scheme to vaccinate girls (and more recently, boys) aged 12-13 years against HPV 6, 11, 16 and 18.[13] The impact on cervical cancer is already beginning to be seen, with oral cancer prevalence in women expected to follow suit in the near future and, in men, to respond similarly approximately 10 years later due to the delayed investment in male HPV vaccination.

Metastasis of oral cancer involves detachment of cancer cells from the primary tumour, proliferation and migration, mainly via the lymphatic rather than haematogenous (blood) or transcoelomic routes.[14] The lungs, bone and liver are the most frequent distant metastasis sites for oral cancer with the lymph nodes being the primary metastatic tumour sites.[15] Lymph node metastases are found in approximately 40% of patients with oral cancer with 15% to 34% of these cases having no clinical manifestation, relying on diagnostic acumen and clinical imaging to locate them.[16, 17] The presence of metastases, which is reflected in the cancer staging at diagnosis, has a profound effect on the patient's prognosis with only 25-40% of patients with metastasis to the lymph nodes surviving for five years and almost 90% of patients with no lymph node metastasis reaching this important milestone.[18, 19, 20]

Time is therefore one of, if not the, most valuable commodity for patients considering that metastasis is found at first inspection in approximately 32.5% [17] of oral cancer cases. While this thesis does not address the public health issues around delayed attendance at a clinic or clinical examination, there remain significant time delays in the post-biopsy timeline which any form of automation could address, but particularly

image assessment AI technologies which can run 24 hours a day, 7 days a week without fatigue. Therefore, automating diagnosis of malignancies from biopsied tissue samples is a logical step towards improving the prognosis of oral cancer sufferers. An important consideration is the ease of implementing such a system into current practice. The final system should therefore work on biopsied thin tissue sections to align with current preparation and storage equipment and procedures.

1.1.1 Current Diagnostic Pathway

A wide variety of imaging techniques are employed throughout the current diagnostic pathway, which is outlined below in Figure 3.[9]

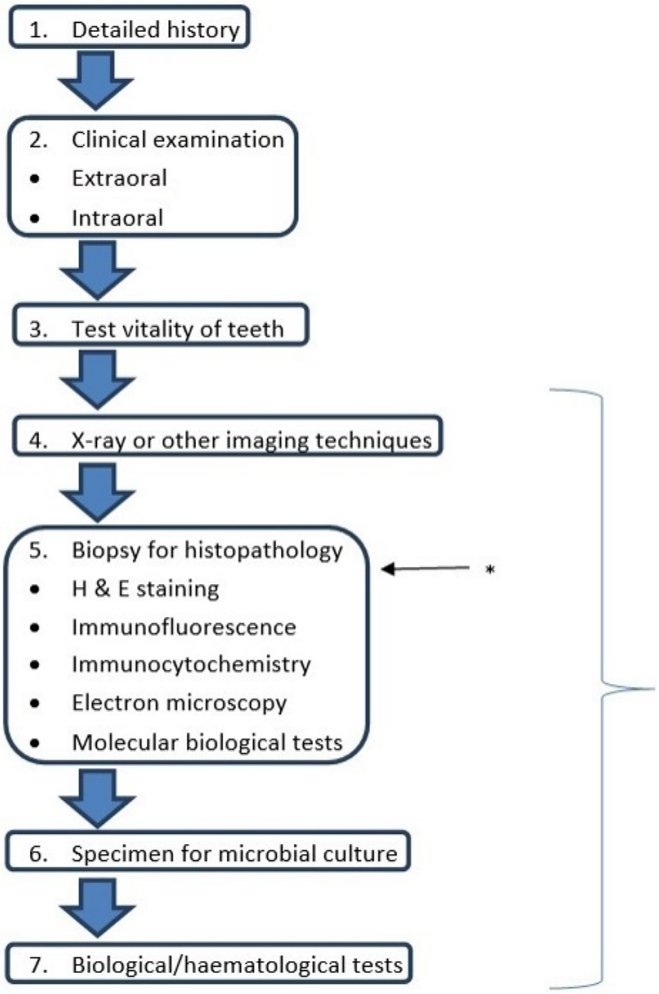

Figure 3: The principles of investigation and diagnosis. Asterisk denotes where our proposed HSI system could be used in place of some additional techniques which are used as appropriate on a case-by-case basis (marked with a bracket). Details of each step are given below.

An oral cancer "check" will begin with discussion of the patient's medical history fol-

lowed by an extra-oral, and then an intra-oral, naked eye resolution inspection. From this stage onwards, a number of imaging techniques are employed for additional diagnostic purposes, assuming enough concern has been raised between the patient's symptoms, signs and history; facilitated by the clinician's interpretation of the constellation of evidential points from the low resolution examination and case history. Not all of these imaging methods (emboldened in the upcoming sections for clarity) will be employed for each patient; rather, specific queries in each diagnosis will dictate the required imaging approaches. Nonetheless, tissue biopsy, staining and inspection by a pathologist is a necessary step in all cases of suspected oral cancer. Surprisingly, as will become clear, this step utilises the least advanced imaging techniques, is entirely qualitative and is highly subjective due to reliance on experienced human interpretation: This is a challenge which spans all medical disciplines managing cases on the data from pathological reports.[21]

Identifying Areas of Abnormality

In assessment terms, the hunt for spread as well as assessment of the primary tumour site is known as "staging the tumour" whose management is critically influenced by the degree of spread beyond the primary tumour site. The following imaging techniques may therefore be employed both in the process of identifying a primary tumour, and in locating potential metastases.

Clinical photography and X-ray imaging are usually the first techniques employed by a dental surgeon as the equipment is available in almost all practices. Clinical photography graphically captures the macroscopic appearance of the lesion at the time, however ongoing comparisons are challenging due to inconsistent lighting, dissimilar cameras, resolution limits and image processing for storage and later display, which can also affect the appearance of the image due to varying monitor settings and viewing environments. The conventional internal imaging available to most general and specialist dental practices is X-ray based, targeted at hard tissue disease such as dental decay or intra-bony lesions which can be developmental, inflammatory, cystic, benign or malignant. Plain dental films (including those with both intra and extra-oral film or detectors), dental panoramic **tomographic instruments** and the most modern "cone beam" CT (computed tomography) instruments are all designed for hard tissue assessment. As the system is optimised for imaging of bone and dental mineral structure, only the lower energy components of the beam are attenuated by soft tissues, leading to a uniform "shadow" contrast only. This minimises any intra-soft tissue shadow contrast, making the systems of little value in cancer diagnostic imaging unless the tumour has invaded mineralised structures such as bone or has originated from within the jaw, similarly destroying mineralised tissues on its way to the surface. If a patient requir-

ing oesophageal cancer assessment cannot tolerate a direct endoscopic assessment, they may have to employ a contrast agent such as barium (administered by swallowing) to assist the clinician in outlining invasive lesions from within the lumen or highlighting exophytic luminal tumours as a contrast filling defect.

If an oral, oropharyngeal or oesophageal tumour is confirmed, **X-ray based imaging** to assess spread elsewhere in the body (most commonly the chest for lung imaging and the neck for lymph node imaging) can be carried out in a hospital to investigate the possibility of distant metastases having occurred. This is vital to asses total patient tumour burden, surgical and overall management decisions, directing intention to cure or if possible, palliative care intentions at the earliest stage.

For further detail of tumour positions in three-dimensions, **computed tomography (CT)** can be employed. In CT scanning, contrast is provided by the attenuation of X-rays (selected for their high penetration depth) as they pass through the body, where they are collected by detectors. The X-ray source and detectors are usually rotated around the body, allowing a 3D rendering of the beam attenuation where contrast encodes information about the scattering and absorption of all tissues, both hard and soft, which the beam passes through. Modified algorithms and use of injected attenuation contrast agents (often iodine based) can further tailor the system's sensitivity to soft tissue anatomy. The diffraction-limited focal spot size that can be produced by X-rays limits spatial resolution of CT scans but, in practice, image reconstruction processes and the cost and design challenges of closely-packed detectors are greater limiting factors.[22] Commonly, a spatial resolution on the order of millimetres is achieved; sufficient to detect areas of abnormality, but orders of magnitude larger than the scale required to identify cancerous cells.

Contrast agents are also used to great effect in **positron emission tomography (PET)** scans, where radiolabelled molecular probes undergo positron decay which interacts with electrons producing gamma photons. This gamma radiation can be used to track the path of the compound they are attached to (typically carbon isotopes in glucose), and therefore probe biological processes. In oral cancer imaging, a radioisotope incorporated into glucose is used to detect metastases (usually in the head, neck or chest) which, like primary cancers, are "hotspots" of metabolic activity due to the fast growing cancer cells. This technique relies on relative contrast with respect to surrounding regions so cancer in areas of high metabolic activity such as the brain, heart and kidneys are difficult to detect with PET scans and if indicated, mandating a separate imaging modality such as MRI. Combined radio and CT detector heads are now used to coregister and record both positron emission and X-ray attenuation by voxel, reporting highly metabolically active tissue in the anatomical CT dataset. Both of these techniques produce substantial levels of ionising radiation which has proven

carcinogenic effects.[23, 24] While the benefits outweigh the risks in some cases, use of these techniques must be limited accordingly.

With no production of ionising radiation, **magnetic resonance imaging (MRI)** records the differential radio emission potential of each imaged voxel, as a result of perverting a uniformly applied magnetic field, aligning all polarising molecules in the body. Energy release upon relaxation of every polarisable molecule provides highly detailed soft tissue contrast data. As bone contains little water and therefore has low proton density, decreasing its polarisability, it appears as a black void defined by surrounding soft tissues and contained liquid marrow spaces.[25] MRI provides the most detailed 3D imaging of a patient's soft tissue anatomy and also clearly highlights tumour invasion of bone. However, its lower spatial resolution, higher running costs and the more complex images it produces are reflected in its less frequent use, relative to PET and CT scanning.[26]

Determining if the Abnormality is Cancer

If X-ray, PET, CT or MRI imaging reveals a head and neck cancer, a sample of the suspicious tissue must be collected by biopsy from the primary cancer and every suspected distant metastatic site in order to make a diagnosis of cancer (or cancer-free). The clinician performing the biopsy is guided by the aforementioned macroscopic imaging techniques and their direct vision alone. As the volume of tissue being excised is frequently smaller than the spatial resolution of the screening techniques, the biopsy and following assessment process are still subject to a range of errors.[27] One study into sources of discrepancy between biopsy and final resection pathology reports found that 60% of discrepancies can be attributed to sampling errors, where the tissue collected is not a representative sample of the suspicious area.[28]

From this stage, there are a wide variety of techniques used to make a diagnosis. The questions which inform the selection of techniques used, and how their outcomes lead to the diagnosis of some of the most common oral malignancies (including cancer) are outlined in the histopathological decisions flowchart below,Figure 4. Evaluations of the mostly commonly used, or broadly applicable, techniques follow.

Figure 4: Example of simplified histopathological decisions flowchart, differentiating multiple benign and malignant pathologies, here using one key criterion for each major decision branch. In reality many factors can influence each decision branch, giving rise to many of the subjective challenges faced by histopathologists [Claire Lock (Viapath) and Professor Richard Cook, private communication] Widefield microscopy of H&E stained tissue sections are used to answer queries in blue.

Almost all oral cancer biopsy tissue samples will be stained with haematoxylin and eosin (H&E) before being inspected, typically using widefield microscopy by an expert histopathologist. Pathology training takes between 12 and 15 years to complete, from medical school through specialty training, and years of additional experience; all of which may be required to identify more subtle indicators of malignancy or pre-malignancy. This is testament to the complexity of biological tissues and cancer pathologies, a fact which is reflected in the difficulties faced by scientists who have attempted to automate this process across the past few decades.[29] Once sampled by biopsy, the soft tissues are preserved with formalin (or live cell transport media if immunofluorescence imaging has been recommended), decalcified and impregnated with wax. This paraffin embedding helps preserve the tissue architecture during sectioning where thin slices of ~4 - 10 µm are taken to minimise scattering contributions from outside the focal plane, negating the need for imaging techniques which are capable of optical sectioning such as confocal microscopy.[30] These thin sections are then mounted on glass slides, stained for tissue contrast enhancement and are preserved under a final glass cover slip. Haematoxylin and eosin are the most commonly used histological stains and, as they are complementary, they are usually used together. Haematoxylin appears blue or violet and binds to DNA, RNA, calcified material, mucopolysaccharide-rich ground substances and some other structures such as granules of keratohyalin. Eosin appears red or pink and binds to the cytoplasm of most cells and of cytoplasmic filaments such as those found in muscle cells and intracellular membranes, keratin, decalcified bone and collagen.[9]

The main indicators of cancer in H&E stained biopsy samples observed with a simple widefield microscope are as follows, in order of increasing severity.[9] (Sub-figures of Figure 5 reproduced, and captions paraphrased from [9] p246-250.)

1. **Blue active nuclei seen in supra-basal cell layers of epithelium**, indicating overactive proliferation and often seen increasingly in progressing severity of dysplasias. (Figure 5 A to B, increasing severity of dysplasia)

2. **Blue nuclei (usually in clusters) seen below the basement membrane** indicating rupture of the basement membrane and epithelial invasion of the internal tissues. (Figure 5 C)

3. **Subsequent intrusion** reflecting increasing local and distant metastatic spread. (Figure 5 C and D)

4. **Specific specialised cell types seen at a distant site**, representing metastasis or invasive spread. (Figure 5D)

5. **The presence of a cell or group of cells that are unidentifiable, or not reflecting their appropriate form at the site in question or a distant metastatic site**. This suggests a loss of architectural and functional cell specificity consistent with their departure from normal biological phenotype control because of genotypic damage: a cancer in progress. (Figure 5 E)

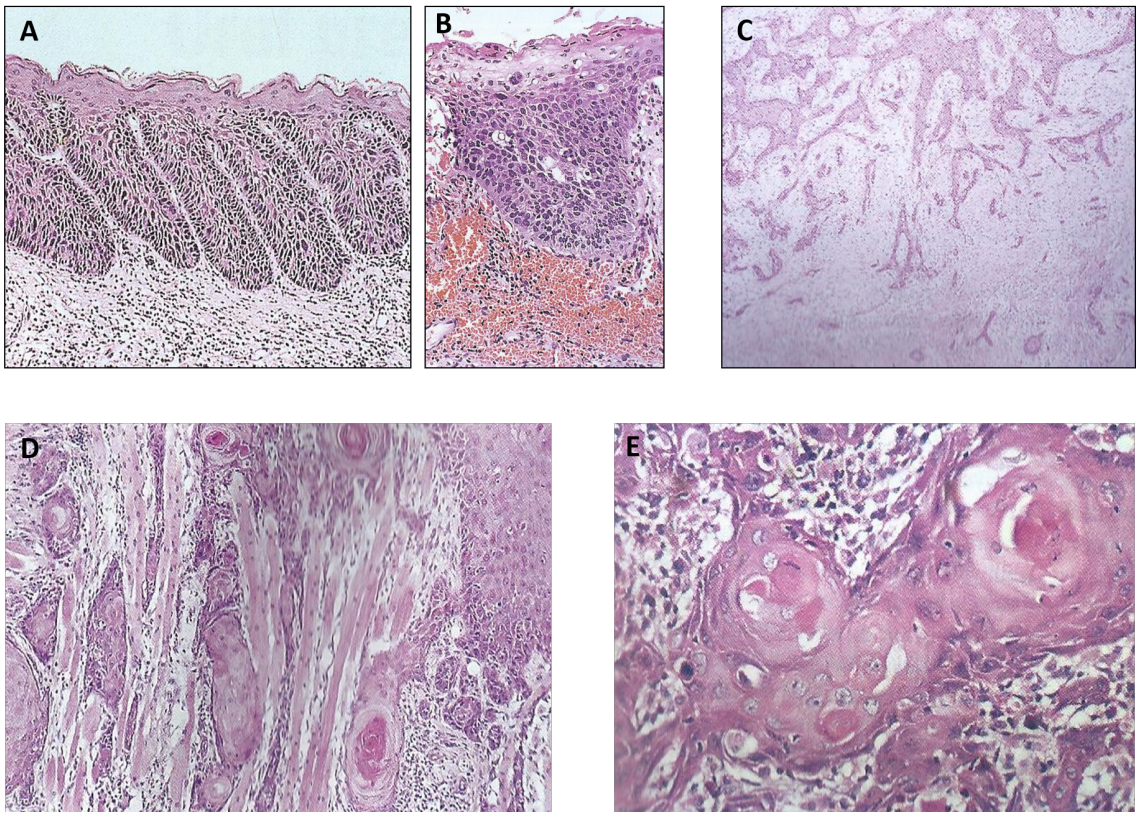

Figure 5: A: Moderate dysplasia. B: Severe dysplasia. C: Squamous carcinoma. D: Squamous carcinoma with cytologic irregularity. E: Squamous carcinoma with keratin pearls around neoplastic epithelial cells. Each intense purple nucleus is on the order of 10 µm in diameter.

It is easy to see how the position of the biopsy sample can significantly impact the reporting of these disease features. For example, a stream of cells invading, having ruptured the basement membrane may be missed by a 4 - 10 µm thick, essentially two dimensional, section of tissue. If the biopsy only catches the inflammatory response of the patient to the invading tissue, the severity and lesion type can be totally misrepresented. To minimise this, many sections are reviewed from each block of paraffin-embedded tissue and in most cases this process is sufficient for an accurate and definitive tissue based, gold standard diagnosis.

The diagnosis itself, from H&E stained sections by a histopathologist is a large source of error in the diagnostic procedure, second only to tissue sampling errors, as indicated

by the large interobserver variation.[28] Kappa values estimating the magnitude of this variability range from ~0.15 to 0.61, tending to be higher for higher grade and no dysplasia and lower for low grade dysplasia. [31, 32, 33] [2]

In some cases, H&E staining is insufficient for making a definitive diagnosis and other stains or imaging techniques (emboldened for clarity) such as immunofluorescence, microbial culture and molecular biology must be employed.

Immunofluorescence uses the specificity of antigen-antibody binding to highlight specific molecules in tissue samples collected by biopsy (immunohistochemistry) or cell samples commonly collected by diagnostic blood smears, fine needle aspirates and swabs (immunocytochemistry). A fluorescent marker is attached to the antibodies (or often they are purchased as a conjugate) and they are then incubated with the tissue section (which must be processed with live cell transport media in place of formalin) to encourage binding to the target molecule.[34] Several different fluorescent markers with different target molecules can be purchased and used in tandem as long as their absorption and emission wavelengths can be separated, usually with filters, to prevent cross-talk between channels and isolation of emission wavelengths. In practice, however, antibody binding site specificity is extremely varied ranging from recognition of enantiomers of the same molecule, to only recognising tertiary structure rather than amino acid sequences.[35] The method can also be inflexible as fluorescence probes that are able to attach to the necessary antibody may be unavailable. In this case, indirect immunofluorescence can be used which involves attaching one or more secondary, labelled antibodies to a primary, unlabelled antibody which is complementary to the antigen of interest. The flexibility of this method is much greater than direct immunofluorescence due to the availability of cheap secondary antibodies that fluoresce at a wide variety of wavelengths. The intensity of the signal they produce can also be increased as more than one (labelled) secondary antibody can bind to the primary antibody.[36] Unfortunately oral tissue is highly autofluorescent, especially at low excitation wavelengths, which can lead to high levels of background fluorescence, obscuring important features.[37] In addition, formalinisation further enhances autofluorescence of structural proteins, in particular, collagen, further aggravating these issues. Immunofluorescence is also negatively impacted by photobleaching which occurs when photochemical alterations to a fluorophore permanently prevent its continued fluorescence. While there is no recognised single antibody marker for oral cancer, immunofluorescence is increasingly used to demonstrate cellular lines of origin in poorly differentiated tissue and as markers for behaviour such as epithelial growth factor receptor fluorescence quantification to assess outcome in oesophageal cancers.[38]

[2] Kappa value is a measure of agreement where 1 is perfect agreement, ~ 0 represents the agreement one would get by chance, and < 0 is weaker agreement than by chance.

In disorders with single diagnostic auto-antibody targets, autofluorescence has become a simple and effective gold standard diagnostic for several oral mucosal diseases including pemphigus vulgaris, paraneoplastic pemphigus, bullous pemphigoid, epidermolysis bullosa acquisita, linear IgA disease, dermatitis herpetiformis, lupus erythematosus and lichen planus.[39] Several have the potential to become malignancies in the long term, or can be a marker for neoplastic disease elsewhere in the body.

Though rarely employed in the oral malignancy diagnostic pathway, the very high resolution of **electron microscopy** (conferred by circumventing the diffraction limitations of light by using an electron source in place of a photon source) can also be exploited to more closely examine ultrastructural cell features that are indicative of epithelial dysplasia. These include ruptured basal lamina and/or hemidesmosomes, pathological cytoplasmic processes and a variety of alterations to the keratinization (such as fewer keratohyaline granules). [40] Whilst light (and occasionally electron) microscopy of most pre-malignant lesions can define and demonstrate epithelial dysplasia, this does not guarantee that the lesion will, or will not, become malignant and some non-dysplastic lesions may become malignant *de facto*, without passing through the phases of dysplasia. Epithelial dysplasia is therefore not a fully reliable marker of malignant transformation potential although, if persisting after correction for all other local driving factors such as alcohol, tobacco and betel quid use, can suggest increased risk of malignant conversion. [41]

In these cases, it is useful to test for the presence of certain biological markers of malignancy. The method by which these molecules are tested for depends on the molecule itself and include immunofluorescence as described above, mass spectroscopy and biological assays; the latter of which is a delicate procedure and therefore the use of molecular biological testing is not widespread. Mitochondrial DNA mutations have been used as disease markers for years and matrix metalloproteinases (MMP) are also commonly used markers. For example, overexpression of the mRNA of MMP-1 and MMP-9 have been linked with oral dysplasia progressing to oral cancer.[42] Growth factors (and their receptors), altered expression of cytokeratin, $\alpha\nu\beta6$ integrin and blood group antigens are more recent additions to the array of molecular indicators used to reveal malignant potential in oral tissue.[43] **Molecular biological testing** encompasses a variety of techniques such as PCR (polymerase chain reaction) for rapidly identifying bacteria and viruses and cytogenetics and in situ hybridisation for detecting genetic abnormalities. For example, tetranectin abundance in saliva can be used to distinguish between primary and metastatic OSCC. These techniques are, however, underused in dentistry unless faced with a very difficult diagnosis. This is largely due to their variable speeds and the expert training or dedicated space in an expertly trained laboratory that is required, contributing to relatively high overall costs. Haematological testing is used

more routinely and is instrumental in diagnosing leukaemias, myelomas, some infections and conditions that are associated with anaemia.[9] **Microbial culture** is rarely used in spite of the infective nature of many oral diseases. Gram-stained smears for bacteria, PAS (periodic acid Schiff) staining for fungal species such as candida and H&E staining of tissue sections are very effective diagnostic tools to determine most infective aetiopathologies. Few oral malignancies are associated with bacterial primary infection, but candida yeast driven hyperplastic mucosal infections and lesions arising from viral infections (particularly herpes virus serotypes HPV-16 and 18) are accepted as drivers of potentially malignant lesions. (This fact was one of the important driving factors in the the National HPV vaccination programme to derive protection for female cervical cancers and penile and oropharyngeal cancer processes in both genders.)[13] The three main occasions where microbial culture is used are to determine which pathogens are present in soft tissue infections to ascertain their sensitivity to antibiotics, to verify the presence of candidiasis (although it is often clear to a histopathologist from H&E stained tissue sections) and to screen for viral antigens or identify a virus using viral culture.[9]

In summary, it is evident that most of these techniques are not broad screening techniques but require prior knowledge or informed assumptions about potential malignancy in order to select which biomarkers to test for. Every technique used increases time, cost, and the potential for error and observer bias. Some techniques may also require additional clinic-pathological information or additional sampling from the patient to correlate clinical symptoms and signs with the tissue structural and behavioural data from the pathology laboratory data. The more stages which require human input, either from the patient, clinicians, pathologists or specialised laboratory staff, the greater the potential for an additional bottleneck in the diagnostic procedure. Also, while the arsenal of additional imaging and testing techniques is ever-growing, the stalwart and compulsory diagnostic step of pathologist-inspected H&E stained tissue sections has remained largely unchanged for decades. In spite of the years of expertise held by all practising histopathologists, there is a great deal of subjectivity and variation in cancer diagnoses which is an unavoidable attribute of the qualitative techniques currently employed. It is clear that the gold standard diagnostic procedure has significant potential for improvement, though it is remarkably successful given the manifold challenges the tissue faces during the interpretation journey.[28] Optimising the accuracy, reliability and speed of this vital step of the diagnostic pathway would, therefore, positively impact the overwhelming majority of oral tissue diagnoses.

1.1.2 Automated Diagnosis

General poor tissue contrast makes label-free imaging of oral cancer very difficult with alternatives which garner additional contrast from another source not performing well, usually in terms of a lack of specificity. For example, autofluorescence is often strong from collagens but poor from other tissues and elastography or similar techniques using tissue texture for contrast struggle to distinguish tumour from scar tissue, amongst other limitations. Therefore, while proteomics, optical techniques such as second harmonic generation (SHG) microscopy and OCT, mass spectrometry and other imaging techniques provide solutions to certain aspects of this vital healthcare challenge, automated diagnosis is a topic of ongoing research.[44, 45] In order to be a reasonable adjunct to, and possible future replacement of, histopathology, the system should be fast, as unsupervised as possible and have high sensitivity and specificity. Whilst a point-of-care, *in vivo* diagnostic device may be the ultimate goal, there are many such high performing, simple tools which have not yet been adopted by the NHS.[46, 47] This reticence may be due to the magnitude of procedural and infrastructural changes that would have to occur to bring these devices into fruitful use.

In order to be easily adopted into the current diagnosis routine, or tested alongside it, the system should comprise a fully automated workflow from tissue section scanning to quantitative diagnostic score. It should also work on tissue sections, primarily as this is one of the most familiar technologies to the current workforce and reduces risk of mis-sampled tisue as is the risk with cyto-smear tests.[48] Another potential reason for these in-clinic devices not being widely utilised is a lack of trust in them. Initially, it is hoped that the proposed quantitative technique could play a supporting role; improving consistency between different histopathologists and providing a quick second "opinion" in complex cases, building trust in their accuracy amongst clinicians.

In the following section we will start by considering, in light of the biological features of oral cancer, the image contrast mechanisms available when light interacts with tissue.

1.2 Light and Tissue

The interaction between light and matter has been a subject of scientific research for hundreds of years, from the understanding of the sky's blue hue and the lenses in the eyes to the inventions of the radio and light microscope. The violet to red portion of light which can be seen with the naked eye is called visible light and is described by photons (single quanta of light) which exhibit a wavelength of ~ 380 - 750 nm, which corresponds to an energy range of 0.51 - 0.26 x10^{-18} m^2 kg s^{-2} by the equation $E = hf$ where h is Planck's constant and f is wave frequency. Photons can interact with matter in two ways; absorption and scattering. Absorption occurs when the energy of the

photon matches the energy gap between two electronic states within the atom causing electronic transitions as the electron absorbs, and then later re-emits, the photon. Light scattering can occur irrespective of the energy of the incident photon. When matter scatters light without fully attenuating it, an interpretable signature of that matter is encoded in that light. This means we can use light scattering as a means to interrogate unknown samples.

Light can be described both by quantum and classical mechanics; by considering light as discrete packets of energy called photons, and as harmonically oscillating electromagnetic fields. The classical behaviour of electromagnetic fields provides more intuitive explanations of the phenomena being discussed in this work, and can be fully explained by Maxwell's equations (Equations 1, 2, 3 and 4) where **D** is electric displacement field, ρ is free electric charge density, **B** is magnetic flux density, **E** is electric field, **H** is magnetic field strength and **J** is free current density.

$$\nabla \cdot \boldsymbol{D} = \rho \tag{1}$$

$$\nabla \cdot \boldsymbol{B} = 0 \tag{2}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \tag{3}$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} \tag{4}$$

Broadly, scattering can be split into two categories: elastic and inelastic, depending on energy transfer during the scattering event. It is possible to use this classical framework to help us understand elastic and inelastic scattering and their dependencies. When light, or an electric field, are incident on a molecule, it induces a dipole moment which has a harmonic time dependence characterised by the frequency of the light and the vibrations of the molecule. The electric properties of a molecule are represented by the polarizability tensor, α. (For the sake of this broad explanation the faster motion caused by the excitation of electrons, and which give rise to resonant scattering, will be ignored.)

The first term of the Taylor expansion of the linear induced dipole moment $\mu^{(1)}$ induced by a electric field $\mathbf{E}_0$, harmonically oscillating with frequency ω_0 on a molecular system can be expressed as follows, in Equation 5 :

$$\mu^{(1)} = \alpha_0 \cdot \boldsymbol{E}_0 cos\omega_0 t + \frac{1}{2}\left(\frac{\partial \alpha_k}{\partial Q_k}\right)_0 \cdot \boldsymbol{E}_0 Q_{k0} cos(\omega_0 t + \omega_k t + \delta_k) \tag{5}$$

$$+\frac{1}{2}\left(\frac{\partial\alpha_k}{\partial Q_k}\right)_0 \cdot \boldsymbol{E}_0 Q_{k0} cos(\omega_0 t - \omega_k t - \delta_k)$$

where α_0 is the first term (at equilibrium position) of the Taylor expansion of the polarizability tensor resulting from the k^{th} vibration, t is the period of the harmonic oscillation, Q_k is a coordinate of normal harmonic vibration at the frequency ω_k, described as $Q_k = Q_{k0} cos(\omega_k t + \delta_k)$.

This expression has three clear terms which give rise to radiation at ω_0, $\omega_0 + \omega_k$ and ω_0- ω_k which account for Rayleigh, anti-Stokes Raman and Stokes Raman scattering respectively.[49]

Rayleigh and Raman scattering are two specific types of scattering which occur when light interacts with tissue under certain conditions. They can each be used to retrieve different information from the tissue according to the mechanism by which they take place, which is our next topic of enquiry.

1.2.1 Inelastic Light Scattering

In inelastic scattering, the energy and direction of the incident particle changes. Inelastic light scattering mechanisms include Brillouin, Compton and Raman scattering where Brillouin and Compton scattering are less probable and occur when photons are scattered by acoustic phonons and charged particles, respectively. Compton scattering requires high energy photons such as X-rays or gamma rays in order to overcome the binding energy of the electrons, and Brillouin scattering involves a change in polarizability of the medium caused by lattice vibrations which are much smaller than the molecular vibrations which Raman utilises.

Raman Scattering

Raman, which is the primary inelastic scattering contrast mechanism in tissue, is when light scattering with a vibrational transition occurs where the molecule changes in polarizability, with the scattered light being emitted in all directions. In the case of inelastic scattering of light by an atom, ground and excited states correspond to electronic energy levels. However, in the case of more complex samples from molecules to dynamic systems like live tissue, the atoms which they are made of rotate and vibrate relative to one another as well as each having distinct electronic energy levels. Raman scattering can take two forms; Stokes and anti-Stokes. As shown in Figure 6, the former is characterised by a decrease in frequency compared to that of the incident light, whereas the latter is characterised by an increased frequency compared to the incident

light. Anti-Stokes Raman scattering therefore depends upon the system already being in an excited state when the light is incident on it, and it is therefore less probable than the Stokes mechanism. Raman scattering has a lower scattering cross sections than elastic scattering mechanisms, such as Rayleigh scattering, which means they are much less likely to occur with just 1 in approximately 10 million photons being Raman scattered. The infrequency of Raman scattering incidents is due to the improbability of a change in molecular polarisation and the energetic difficulty of exciting low frequency vibrations with a high frequency radiation.[50]

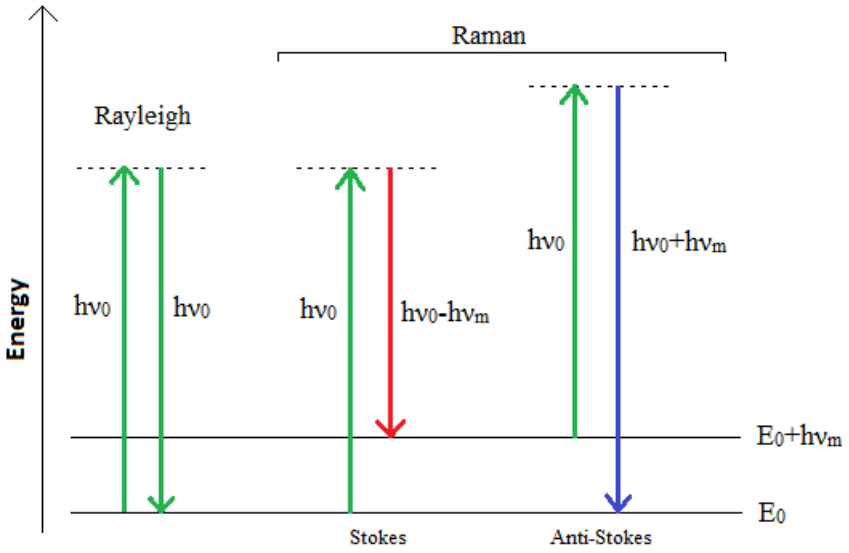

Figure 6: Jablonski diagram depicting energetic transitions resulting in Rayleigh Stokes Raman and Anti-Stokes Raman scattering where $\nu_0$is the wavenumber of the illumination and ν_m is the wavenumber of the vibrational transition.

As Raman scattering is inelastic and therefore involves some energy transfer which relates to the atomic structure of the constituents of a sample, the resulting light contains more detailed, chemically specific information about it. Raman spectroscopy provides, without requiring the use of dyes or external contrast agents, sample-specific information about the molecular composition and interactions in the form of a vibrational spectrum. As this spectrum is unique to each sample, it is often called a "molecular fingerprint".[51] Raman spectroscopy helps to identify the molecular structure of materials and is often used to help identify or characterise samples of unknown composition. This has been exploited to great effect for probing biomolecules for over 30 years, and specifically for investigating cancer with an aim to diagnose it for over 20 years, starting with the work of Gniadecka *et al* on basal cell carcinoma.[52][53]

Segmentation of Raman spatial maps of tissue areas show good agreement with

histopathology and therefore the source of this contrast is a popular line of enquiry. While there are differences in the Raman signatures of different kinds of cancer, a general increase in the intensity of spectral bands associated with DNA and protein are often observed in cancer tissue.[54] Biomolecules which can be used to help map cancer are still being uncovered. Keratin and water have been identified as significant contributors to the discriminatory components of oral Raman spectra. In the first instance, Chen *et al* developed a new analysis protocol based on spectral unmixing of the molecular fingerprint of the well-known cancer biomolecule, keratin, which could detect cancerous tissue with 100% specificity.[55] This is a good example of Raman spectroscopy's capacity for very high specificity due to its chemical origins. Secondly, Barroso *et al* used the ratio of the Raman bands at 3,390 cm^{-1} and 2,935 cm^{-1} to calculate water concentration in human oral squamous cell carcinoma (OSCC) samples. Significant differences were found, with a decrease from $54\% \pm 24\%$ to $76\% \pm 8\%$ across the healthy tissue-tumour border. This was used very successfully to delineate between healthy tissue and OSCC with a 98% accuracy. In subsequent work it was shown that the distance from the tumour border at which water concentration, and its standard deviation, begin to change is remarkably consistent, at ~1.5 mm. Identification of adequate tumour resection margins using this principle was evidenced.[56] In work by Behl *et al*, inflammatory cells were successfully differentiated from tumour cells, a feat which would be difficult to achieve with no chemical information. Morphologically they can appear similar due to the enlarged nuclei often present in both cell types and the pattern of inflammation, which tends to closely precede the leading edge of the cancer. It may also be of particular note in oral tissues where it is especially challenging to separate inflammatory conditions such as lichen planus from pre-malignancies such as leukoplakia.[57]

In spite of the very high specificity, high sensitivity to many relevant biomolecules and the potential for unsupervised analysis which make Raman spectroscopy a highly suitable technique to support or mimic histopathology, it has several characteristics which prevent it from being the ideal technique for this purpose. Firstly, it is usually performed on fresh/frozen tissue as the strong signal (both Raman and autofluorescence) from traditional fixatives obscures important spectral information. Cluster analysis methods improve base spectrum isolation and may help to alleviate the problem of fixative signal but, as with many post-processing techniques, it comes at a cost to the SNR. This can be problematic as biological signatures are often less intense and spread across "fingerprint regions" which are unique to different compounds but which contain many peaks, of which very few are individually identifiable, requiring more complex identification processes. It does mean, however, that in this respect Raman spectroscopy is well suited to *in vivo* application and therefore the development of "opti-

cal biopsy" systems to prevent destructive cancer sampling for diagnosis is popular.[58] Unfortunately as Raman scattering is improbable and slow it cannot be used to map areas *in vivo*, only to collect point spectra. As the number of Raman-scattered photons is inversely proportional to the wavelength of the incident light, theoretically the strength of the Raman signal could be increased by illuminating the sample with shorter wavelengths. However, in practice this is not possible due to the strong fluorescence of biological samples in this wavelength range which can mask the weak Raman signal. Therefore, NIR illumination is preferred for tissue imaging as a compromise between the undesirable fluorescence at lower wavelengths and the desirable increased Raman signal yield. Coherent Raman mapping can be performed more quickly, and to a higher penetration depth, due to the signal enhancement conferred by the coherence of the Raman scattered light (and with decreased potential for fluorescence interference due to signal collection at the blue end of the visible spectrum).[59] However, as a third order nonlinear optical technique, it requires two coherent pulsed laser sources to drive the in-phase molecular vibrations and it is therefore significantly more expensive to build and run than spontaneous Raman spectroscopy systems which require just one continuous wave laser.[60] The additive nature of the coherent signals also biases the component reporting towards the most abundant Raman-active molecule in the illumination spot which could result in inaccurate segmentation maps.[61]

One idea to economically overcome this speed limitation is combining Raman spectroscopy with a fast, widefield optical technique for use in a pre-screening protocol to identify regions of interest (ROIs).[62] Vibrational spectra acquisition could then be limited to the coordinates of these ROIs. A range of elastic light scattering based optical techniques were considered for this application.

1.2.2 Elastic Light Scattering

In elastic scattering, a photon interacts with matter but there is no energy transfer between the two. Gustav Mie was the first scientist to calculate Maxwell's equations for specific systems, and he did so first for elastic light scattering where the spherical scattering particles are similar in size to the incident illumination light wavelength. From this solution, he was also able to use Mie theory to model two other elastic light scattering domains, covering a large proportion of commonly observed electromagnetic radiation scattering effects. As a result, there are three main models to describe elastic light scattering which depend on the size of the particles (πd) and the wavelength of the incident light (λ) which can be combined to create an arbitrarily defined size parameter, $x = \pi d/\lambda$. As well as the relative sizes (and shapes) of scattering particles and the incident light wavelength, the amount and direction of scattering depends on many

other factors including light polarisation and coherence, and its angle of incidence.[63] There are three sets of constraints, defined in terms of x, in which different assumptions can be applied to the Mie scattering model.

The most complex of the three systems occurs when roughly spherical scattering particles are illuminated by light with a wavelength similar to the particle size, x $\approx$ 1, and is described by the Mie scattering regime. [64] Mie scattering occurs in all directions, with a stronger forward-scattering component which increases with particle size. The theoretical scattering cross section, σ_s, in the Mie regime is described by Equation 6.

$$\sigma_s = k^{-2} \int_0^{2\pi} \int_{\pi}^{0} (\mid S_1 + S_4 \mid^2 + \mid S_1 + S_4 \mid^2) dcos\theta d\phi \tag{6}$$

where S1 and S4 are terms from the scattering amplitude matrix (which must be calculated from Maxwell's equations for the boundary conditions of the scattering event), d is the scatterer diameter (assumed spherical), k is the wavenumber, and θ and φ are angles which describe the propagation direction of the scattered light $\boldsymbol{r} = \boldsymbol{r}(\theta, \varphi)$.[65] It is not trivial to solve Maxwell's equations explicitly, which prohibits Mie scattering solutions from being used for much more than examples of the behaviour of scatterers of certain shapes or composition. A subset of Mie scattering, called Rayleigh scattering, occurs when the scattering particles are much smaller than (at most 1/10 of) the wavelength of incident light, x $\ll$ 1. This is considered to be a subset of Mie scattering as, due to their small size, even non-spherical particles can be modelled as spheres of equivalent volume. The small size of the scatterer relative to the incident light paints a mathematically relevant picture of the particle being "submerged" in homogeneous electric field. In a Rayleigh scattering event, a photon of a certain wavelength is absorbed by exciting bound electrons to higher energy states, and another photon of the same wavelength is emitted simultaneously and isotropically (conserving energy), like a dipole. The isotropic polarizability α can be used to simplify the scattering amplitude matrix, reducing the scattering cross section to a much simpler form, shown below in Equation 7:

$$\sigma_s = \frac{8}{3} \pi k^4 \alpha^2 \tag{7}$$

As wavelength is the reciprocal of wavenumber, this equation demonstrates the characteristic λ^{-4} wavelength dependence which broadly defines Rayleigh scattering. We can see this relation in action when the sky appears blue on a clear day where the shorter (blue) wavelengths of the white light from the sun are more strongly scattered by small atmospheric particles than longer wavelength.

When the scattering particles are much larger than the wavelength of incident light, $x \gg 1$, geometric scattering occurs. When the scattering particle diameter is more than ten times the wavelength of incident light, whilst Mie theory can be used to describe the scattering outcomes, it is more common to turn to the simpler laws of geometric optics to describe the system. In any case, geometric scattering is relatively minimal in the case of tissue sections, which are rich in features much smaller than the micron (and above) scale scatterers required to produce geometric scattering of visible light.

As elastic scattering does not involve any energy transfer which relates to the atomic structure of the constituents of a sample, the information encoded in the scattered light contains less specific, chemical information than light which has been inelastically scattered. However, the higher scattering cross section means that elastic light scattering comprise the primary mechanisms by which light and tissue interact.

Tissue is an ideal scattering medium, with most comprising structures with a wide range of sizes from a few tenths of nanometres to a few tenths of micrometres.[66] They can be described as a random continuum of the inhomogeneities of the refractive index with a varying spatial scale.[67, 68] A wide range of intracellular structures affect tissue light scattering. The diameters of cell nuclei are on the order of 5-10 µm, mitochondria ~1-2 µm, lysosomes and peroxisomes ~20 nm with structures within organelles ranging in size from a few to a few hundred nanometres. Many cells and subcellular components are approximately spherical or ellipsoidal and are therefore often modelled as collections of homogeneous spherical particles which can be well described by Mie theory.

Previously in the introduction, the H&E-visible features which are the foundation of the cancer grading system used by histopathologists were outlined (1.1.1). These features, and other markers of cancer such as enlarged nuclei, can be "translated" into scattering effects whose identification can be automated. For example, measurements of refractive index across a range of wavelengths could be used to theoretically recover the distribution of nuclear sizes in a sample for use in a diagnostic capacity.

The refractive index variation within tissues can be quantified by the ratio between the mean refractive indices in the sample's scattering centres and the refractive index of its surroundings. This ratio also impacts the efficiency of the light scattering that occurs. Squamous cell carcinoma is a common and potentially devastating oral malignancy which begins in the squamous epithelium. Nuclear-to-cytoplasmic ratio, and nuclear size have both been observed to increase in cancerous tissue, increasing three to five-fold from ~ 4 - 7 µm in diameter to ~ 20 µm.[69] Epithelial nuclei can be modelled as spheroidal Mie scatterers with a higher refractive index than surrounding cytoplasm. Considering nuclei as optically soft [3] and assuming illumination with visible

[3] Optical softness is when the refractive index of a particle is similar to the refractive index of the media around it.

wavelength light which is significantly smaller than the scatterer sizes ($\lambda << d$), the van de Hulst approximation for "large" spherical particles can be applied to describe the elastic scattering cross section of the nuclei as follows, in Equation 8 [70]:

$$\sigma_s(\lambda, d) = \frac{1}{2}\pi d^2 \left[1 - \frac{2sin\delta}{\delta} + \left(\frac{2sin\delta}{\delta}\right)^2\right] \tag{8}$$

where $\delta = \frac{2\pi d(n_{nucleus} - n_{cytoplasm})}{\lambda_0}$, d is the particle diameter, n is refractive index and λ_0 is the wavelength of light in a vacuum.

The scattering cross section, therefore, varies periodically with inverse wavelength which also generates a periodic component in tissue optical reflectance or transmission. Fourier space can be used to probe frequency (inverse wavelength) and, in combination with measurements of refractive index across a range of wavelengths, could be used to recover the distribution of nuclear sizes in a sample for use in a diagnostic capacity.

This general framework is used frequently in elastic scattering probing of cells and tissues: mathematical modelling of tissue scattering is developed to describe a system, the model is varied iteratively until a best fit to the data is achieved and then a known relationship between a parameter from this model and the physical measurement is used to estimate said parameter. In reality, the accuracy of both the chosen model and the selection of "known" values used to constrain the iterative fitting of the model can vary hugely. This is especially problematic when comparing *in vivo* and *in vitro* measurements which differ largely due to differing hydration levels. With the large number of assumptions made by this general approach, including scatterer size, shape and refractive index and tissue morphology and composition, the cumulative errors can be large.[71]

Light scattering spectroscopy (LSS)

Light scattering spectroscopy (LSS) uses spectra collected from singly scattered photons to probe tissue morphology, by analysing it with Mie theory or more sophisticated models, as in the approach outlined above. By creating systems or samples which make multiple scattering events unlikely, it is possible to attribute changes between properties of the incident and resultant wave at each point on the detector, to the sample at the corresponding position. Multiple scattering events add confounding factors, both in terms of the position of the scatterer, and its properties. The effects of each scattering incident will cumulatively affect the incident light, and cannot be uncoupled from each other with the information available at the point of detection. There are many means by which scattering is constrained to single events, such as by using illumination wavelengths with low penetration depths into the medium being imaged, or by creating very thin samples.

As previously mentioned, cell nuclei can give rise to Mie-type scattering which is visible in spectra as a broad sinusoidal curve. As elastic scattering spectra are usually comprised of various scattering features (Mie and Rayleigh) and absorption bands (more common *in vivo*), this slowly fluctuating baseline can affect both the wavelength and the intensity of other spectral features, often broadly called "dispersion artefacts". More recent work has shown that these artefacts can be largely explained by resonant Mie scattering.[72] When resonant Mie scattering occurs, there are several deviations from basic Mie theory, such as the refractive index becoming wavelength dependent near absorption bands, requiring application of the Kramers-Kronig relationship.[73] Additionally, Lau *et al* have also shown that the regular packing of haemoglobin in blood vessels is a substantial confounding factor when it comes to model-based LSS and ESS (elastic scattering spectroscopy) tools, especially *in vivo* or in bulk tissue *ex vivo*.[74]

This diagnostic technique depends on having accurate models to apply to real life data for extraction of values for parameters such as reduced scattering coefficient (μ_s'). Across the visible and infra-red wavelength ranges, recovery of the "pure" spectra (absorption or an individual scattering type such as non-resonant Mie scattering) required for this is difficult, even if the model and parameters selected for this analysis are a good fit for the real data. This inhibits reliable comparison across samples.[75] With our understanding of elastic scattering in complex media changing so rapidly, a diagnostic method which utilises scattering signals without requiring a complete appreciation of the origins and behaviours of scattered light in tissue would be prudent.

Elastic scattering spectroscopy (ESS)

Elastic scattering spectroscopy (ESS), or diffuse reflectance spectroscopy, utilises the spectra of light which has been diffusely scattered by tissue to retrieve tissue optical properties for diagnostic purposes. While ESS spectra can, like in LSS, be analysed by fitting them to a model, the diffuse light scattering spectra produced are significantly more complex which makes this approach less appealing and less common.[76] Additionally, as light which has undergone multiple scattering events can be included, ESS spectra tend to be simpler to collect, lending the technique to *in vivo* studies. For both of these reasons, a broader range of potential analysis techniques are apparent in the literature including direct analysis of specific wavelength regions of the spectra and unsupervised clustering such as by LDA.[77] First and second order derivatives of diffuse reflectance spectra (DRS), collected at a range of distances by a handheld probe, were used to develop a classification algorithm in work by de Konig *et al.*[78] This technique was designed for *ex vivo* intraoperative use, to evaluate the deep resection plane of oral cancers. Many DRS techniques utilise the distribution of coloured molecules in tissue,

such as haemoglobin to create diagnostic metrics.[79] In this intraoperative application, the presence of blood which is not contained within vasculature prevented use of this common approach and necessitated extension of illumination wavelength range into the near-infra-red (NIR). Diagnostic sensitivity and specificity using a linear support vector machine in a five-fold cross validation process depended on measurement probe distance but were as high as 82% and 89%, respectively.

In later work with the same system, it is referred to as near-infrared hyperspectral imaging rather than diffuse reflectance spectroscopy.[80] This belies the large overlap between the two techniques.

Hyperspectral Imaging (HSI)

HSI is a broad term referring to techniques which collect the transmission, fluorescence, reflectance or scattering (elastic or inelastic) by a sample of different, discrete wavelengths of incident light. An intensity against wavelength spectrum for is constructed from this information for each pixel or sample region. These spectra are unique to different materials which can be used to help identify different components of the sample. Whilst it can be said that Raman spectroscopy and ESS, for example, are hyperspectral techniques, they would rarely be called HSI. When a more specific definition is available for a technique, "hyperspectral" is rarely used to name it. Most modes of HSI can be performed widefield, preventing damage to delicate tissue samples with low illumination intensity and providing high imaging speeds making it ideally suited for tissue screening.

There are many ways to illuminate each "pixel" of a sample at multiple wavelengths, and the technique chosen depends on the application and the available equipment. There are four modes; spectral, spatial, spatiospectral and non-scanning, each with several different scanning techniques. **Spectral** scanning involves taking a 2D (x, y) image of a sample at every wavelength (λ) within a range to create a "hypercube", as shown in Figure 7.

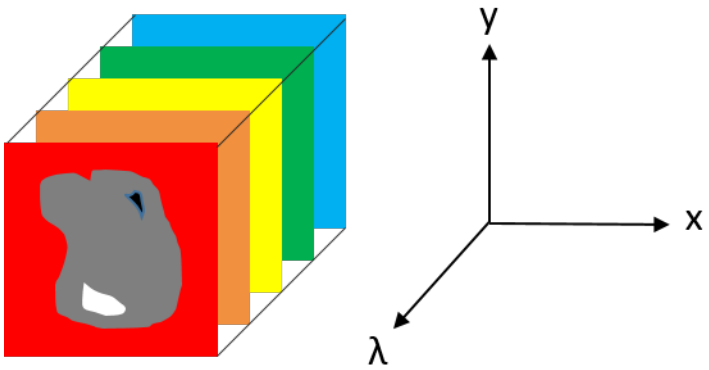

Figure 7: A hypercube, with different illumination wavelengths (λ) shown as coloured fields of view (x,y), stacked into a "cube" of data.

A tuneable light source illuminates the sample at one wavelength and takes a wide-field image. The system then continues to alternate between taking an image and changing the wavelength until the full wavelength range has been covered. The step sizes should be as small as possible (to produce the highest spectral resolution) and found by iteratively decreasing the step size until no additional features appear [81] or the bandwidth of the light source becomes the limiting factor. The spectral and spatial data are then combined to make a 3D (x,y,λ) data "hypercube".

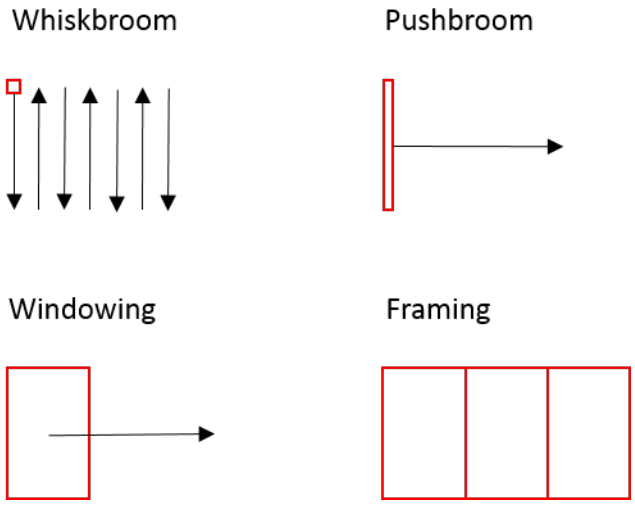

Figure 8: The four main hyperspectral imaging spatial scanning modes.

In **spatial** scanning hyperspectral techniques, a full slit spectrum is produced for each pixel or row of pixels by one of four main modes shown in Figure 8. The basic setup is comprised of a prism-grating-prism optical system (or similar) with a sample placed on a moving stage so that pointwise or linewise spectra can be collected by a camera.[82] These point or line scanning techniques are known as whiskbroom and

pushbroom, respectively. Two rarer scanning techniques are windowing and framing. The former involves the scanning of a 2D (x, y) frame and the latter creates a grid of 2D frames, each one field of view (FOV), that cover the whole sample. [83]

Spatiospectral scanning is the most recently developed mode and it produces diagonal slices of the hypercube, each of which is a spatial map of the scene in every wavelength of the chosen range.[84] The technique is difficult and its theory was proposed long before it was practically achieved.[85] These setups include a slit with a dispersive element onto which projections from a pinhole array are projected, and a camera that adds a 2D (x, λ) strip to the hypercube which is completed by scanning the camera perpendicular to the slit.

Snapshot and **staring** hyperspectral imaging are non-scanning modes. Snapshot HSI systems have no moving parts and instead record the full spatial and spectral details of a field of view with each frame. As the system does not have to scan in any dimension, it can be used at video-rate to monitor moving objects and active processes.[86] The wavelength is changed continuously and the full hypercube is recorded in one frame with no moving parts or scanning required.[87] There are now several different methods by which snapshot systems work including fibre-reformatting imaging spectrometry, computed tomographic imaging spectrometry and image mapping spectrometry.

Spatial scanning is used most commonly in remote sensing and surveillance, spectral scanning in microscopy applications where one field of view can cover the majority of the sample and spatiospectral has potential in monitoring active processes.

HSI Analysis

The data produced by all hyperspectral methods are 3D (x, y, λ) hypercubes.(Figure 7) With such large volumes of high-dimensionality data, analysis options are abundant. Samples produce unique spectra depending on how they interact with photons of different wavelengths. In most fields utilising HSI, samples are coloured and therefore most spectral features are from absorption. As a result, a large proportion of HSI analysis work relies heavily on "matched filtering" (a type of partial unmixing) where collected spectra are matched with reference spectra from a relevant library in order to identify sample components.[88] Mie scattering by nuclei and mitochondria is the greatest contribution to visible and NIR scattering spectra in tissue so HSI performed on unstained biological samples will produce spectra mostly comprised of Mie scattering, some Rayleigh scattering and occasional absorption from residual pigments such as haemoglobin. As absorption is minimal and the main source of contrast will be scattering, resulting spectra often lack distinctive features. It is therefore necessary to characterise sample regions by the overall shape of their spectra rather than by comparing specific features from reference spectra, which is more time consuming. The

reference spectrum attributed to a sample region will be the one with the most similar absorption at the most wavelength values. Even with this constraint, analysis relying on matched filtering has two major shortfalls for this application: Firstly, the spectral libraries available are largely for agricultural and surveillance application with no comprehensive biomedical libraries available. Secondly, most systems require the user to input control spectra of the expected image constituents meaning that prior knowledge of the image subject is required which is undesirable for our application which aims to be as automatic as possible with minimal human input. Also naturally occurring materials, such as those we will be investigating, vary in composition and therefore have variable absorption spectra making their identification more difficult than synthetic materials which tend to be more consistent.

K-means clustering achieves a similar outcome, but without the use of reference spectra. This allows spectra to be separated into a user-defined number of distinct groups according to how different they are, without defining the source or "seed" of each spectral group. It has been used to great effect in separating breast ductal carcinoma from healthy breast tissue. Hypercubes of unstained tissue sections were collected with a snapshot HSI system and compared to H&E stained tissue sections segmented by histopathologists. The k-means-based classification achieved a detection sensitivity of 85.45%, and specificity of 94.64%.[89] This technique is often described as semi-supervised as it only requires the number of sample components to be selected by the user. A semi-supervised technique allows us to begin answering the question "Can we segment healthy and/or cancerous tissue in a useful way using HSI?" without making complex assumptions about the data which can be a source of errors, as in the modelling based LSS and ESS approaches discussed previously. As we were not aiming to make a diagnosis based on a specific biological or cancer feature or product, this was sufficient.

HSI data has also been successfully classified using supervised machine learning methods based on multilayer perceptrons and other neural networks. When the data has a large number of wavelength steps compared to number of spectra collected, however, these methods suffer from the Hughes phenomenon where the accuracy of the classification begins to decrease at a certain, high band number.[90] As a result of this overtraining effect, the classifier can have poor generalization capabilities. A classification method widely used in HSI that overcomes the Hughes phenomenon is the use of support vector machines (SVM).[91, 92, 93, 94, 95, 96, 97, 98, 99, 100] Support vector machine (SVM) based methods are popular due to their good generalization capabilities, ability to classify globally and lack of sensitivity to the number of dimensions present.[101] These methods do have some limitations, however: Oral tissue is uniquely complex exhibiting substantial interpatient heterogeneity.[102] The complex anatomy of the oral cavity with differences in histology, keratin content, and diverse lesions

makes it distinctive from other organs. Large intra-class variation makes oral cancer particularly difficult to train diagnostic classifiers for. Machine learning approaches also require access to a larger number of representative samples to be used as training data, and a large time investment from one or multiple histopathologists who would need to classify each and every case.

In neoplastic tongue tissue feature extraction work by Fei *et al*, nucleus and cytoplasm texture features extracted from the epithelium were regularly selected as the features with the greatest discriminatory power.[102] Whilst many studies ignore the spatial content of hypercubes, neighbouring pixels are often highly correlated, and combining this information with spectral classifications can provide additional information.[103]

Texture analysis characterises image regions by looking at the spatial variation of pixel intensity values. As most samples will contain areas of different textures, locating and demarcating the boundaries between these textures can provide useful segmentation. Popular statistical methods for texture examination include first order statistics such as standard deviation of grayscale value, Gabor filters, fractal textures and second order statistics derived from the grey-level co-occurrence matrix (GLCM), defined by Haralick *et al* in 1973.[104]

Texture analysis has been used successfully on a wide range of image types to elucidate cancerous cell and tissue alterations. Chalut *et al* induced apoptosis in MCF7 cells using two chemotherapy drugs; paclitaxel and doxorubicin. By comparing scattering images and fluorescent images at different timepoints they were able to link the increase in the fractal dimension, which correlates with subcellular texture, with mitochondrial and nuclear alterations due to induced death. [4] Understanding apoptosis is seen as a key step in understanding cancer biology and these results. The standard deviations of elastic scattering spectra of tumors are often shown to be much higher than normal tissues and, in spite of being a diffraction limited technique it is sensitive to spatial variations in refractive index below the diffraction limit.[92, 106]

Texture analysis has also provided some potential mechanisms which can help to account for the increased scattering variance in cancer. Modelling of light scattering by cervical cells with varying degrees of neoplasia by Drezek *et al* using a sophisticated finite-difference time-domain model which allows Maxwell's equations to be solved numerically, showed that healthy nuclei are significantly more homogeneous than those in neoplastic cells due to chromatin clumping which occurs in high grade nuclei. They also demonstrated an increase in scattering intensity across all detection angles with neoplasia, in spite of nuclei being predominantly forward-scattering, and an increase in

[4] Fractal dimension is a statistical measure of textural complexity which probes how the features change with scale.[105]

scattering cross section and its variance belying an increase in the nucleus size and DNA content as well as in range of cell size and shape as expected due to pleomorphism.[107]

Existing Work on HSI for Cancer Detection

HSI has been used for many years in the military for surveillance, in agriculture, geoscience and astronomy however its use in the biomedical sciences is more recent and less well explored, particularly in unstained tissue. Work has rapidly progressed to its diagnostic application *in vivo* however realising the use of point-of-care diagnosis techniques in clinics is difficult. As a very fast technique with high sensitivity and spatial resolution, its potential for application in a supporting role in histopathology is clear, and evidenced in the selected works below.

There is little published work on using visible wavelength HSI to image unstained tissue sections although individual cervical cells have been characterised as cancerous, pre-cancerous and benign with success rates of 95.8%, 66.7% and 93.5% respectively using wavelengths 400-1000 nm.[108] Both spectral and spatial information such as nucleus size were used to make diagnoses. Infra-red (IR) wavelengths have been used more successfully. Several groups have used the fast tuning benefits of quantum cascade lasers (QCLs) to this end: Kroger-Lui *et al* rapidly identified goblet cells in unstained, paraffinised thin sections of mouse colon with mid-infra-red (MIR).[109] The spectra were processed by training a random decision forest classifier on one sample with a k-means clustering basis, after which it was successfully applied on five more tissue sections for blind validation. The ratio of absorption at 9294 and 8467 nm and the absorption integral over the wavelength range 9200 to 9737 nm were used in combination to identify tissue components, similar to the matched filtering often performed on non-biological hyperspectral data. Kroger also led another team in using a similar system with a microbolometer array detector to successfully image unstained mouse jejunum with diffraction limited spatial resolution using MIR illumination.[110]

Successful works using HSI for the diagnosis of oral and head and neck cancer (HNC) have almost exclusively been on stained tissue where absorption is the primary source of contrast. Research on unstained tissue, inevitably using scattering as the main contrast mechanism, is less common and analysis of the data in these studies varies greatly. Ou-Yang *et al* demonstrated a sensitivity of 90 $\pm$ 4.53 % and a specificity of 87.8 $\pm$ 5.21 % in the detection of cancer in stained oral carcinoma biopsies.[111] They used an embedded relay lens microscopy hyperspectral imaging system (ERL-MHSI) in transmission and two fluorescence modes, 330-385 nm and 470-490 nm, which benefited from combining morphological and spectral information. The morphological data was processed using fractal dimensions and five different spectral data analysis methods (selected on a case by case basis depending on a number of factors such as patient age).

They used 58 training samples and just 10 test samples. Akbari *et al* reported very good sensitivity and specificity of HNC metastasis detection of 92.6% and 97.7% respectively in lung histological slides and 96.2% and 98.3% respectively in lymph node slides by imaging H & E stained histopathological slides with a hyperspectral microscopy system utilising wavelengths in the range of 450-950 nm.[94] The data were classified using a support vector machine after using 12 slides for training and evaluation by leave-one-out cross-validation. Whilst the cancer cells themselves started life in the oral cavity, the surrounding tissue morphology is extremely different, making the generation of machine-learning based discriminants possible with less risk of overtraining.

Histopathological diagnosis and grading of oral epithelial dysplasia is informed by both architectural and cytological changes. As a result, focus is shifting to imaging modalities and analysis methods which can extract information about both facets of the diagnostic picture; chemical and morphological, such as in the cervical cell work by Siddiqi *et al* outlined previously. Most notably, Fei *et al* used colour, texture, morphological and topological analysis of RGB images of H&E stained mouse tongue sections to extract key diagnostic features in identifying neoplasia. These features were combined to create a predictive model for diagnosis which was evaluated on human tongue tissue by leave-one-out cross validation, yielding a sensitivity and specificity of 100% and 82.7%. Hypercubes of the same mouse tongue tissue, *in vivo* and *ex vivo* (fresh), were also collected across a 450 – 900 nm wavelength range. The magnitude of the Spearman's rank correlation coefficients (r_s)[5] for these diffuse scattering spectra and the key diagnostic features extracted from the histological features were evaluated. In 8/9 features, $r_s \geq 0.5$, though overall correlation was slightly lower in the *ex vivo* tissue than *in vivo*.[102] This bodes extremely well for the interrogative and diagnostic capacity of unstained hyperspectral imaging of oral cancer as a standalone technique.

As HSI has been used successfully *in vivo*, there appears to be great potential for HSI as a fast pre-screening technique and perhaps even a diagnostic tool.[112] Due to the required remodelling of the current diagnostic pathway, point-of-care diagnostic devices are unlikely to be adopted by healthcare providers, evidenced in part by new incentives to support innovative medical device uptake.[113] Efficacy of *in vivo* hyperspectral techniques based on diffuse elastic scattering is affected by the various pigments which occur naturally in the body, such as melanin and haemoglobin. In HSI, highly absorbant haemoglobin is such a strong spectral feature that diffuse scattering information becomes difficult to retrieve, limiting its useful application to instances where the haemoglobin distribution and oxygenation itself can be used as the primary diagnos-

[5]Spearman's rank correlation coefficient is a measure of the strength and direction of correlation between two sets of data. It can take values between -1 and +1 with 0 indicating no association between the two datasets and +1 indicating perfect positive correlation.

tic feature. Examples include *ex vivo*, whole fresh excised tissue samples which retain their blood distribution or *in vivo* applications such as intraoperative tumor margin detection.[114, 115] Additionally, *in vivo* techniques are sensitive to movement and face the challenge of coregistration with H&E stained sections for validation; a step which is non-trivial even in *ex vivo* studies. The difficulty of proving high sensitivity and specificity by this validation process may in part explain low clinical uptake of such devices as it is an inevitable component of the 501(k) pre-market approval process which brings most medical devices to market. Nonetheless, given the abundance of simple, in-vivo, haemoglobin absorption-based techniques available and the minimal uptake in clinic, it seems prudent to focus on developing techniques which dovetail more neatly with the current diagnostic pathway.

There is evidence to give hope that the proposed hyperspectral elastic scattering prescreening modality may be sufficient to segment and diagnose oral cancer, though in histologically prepared tissue samples, and in widefield transmission mode for maximum signal intensity and imaging speed, scattering can be difficult to spatially register as it may undergo several scattering events. Spatial coregistration with an inelastic light scattering technique could help in making diagnostically relevant tissue maps, adding specificity and meaning to the diffuse light scattering. With both elastic and inelastic scattering-based tissue imaging having major limitations for diagnostic applications, a combined system exploiting each technique for its best attributes, namely speed and specificity, is evidently worthy of investigation.[62]

1.3 Summary

Oral cancer is a globally worsening issue with poor survival rates linked strongly to late diagnosis. The current diagnostic pathway involves tissue biopsy and diagnosis by a histopathologist; a process which is qualitative and which therefore has poor repeatability. Hyperspectral elastic scattering microscopy and Raman spectroscopy have been combined very few times before, and never before on unstained tissue sections which would be most compatible with the current oral cancer diagnostic process. The main example of a combined HSI-Raman system comes from Smith *et al* who built a system able to perform both imaging modalities simultaneously using a < 10 µm-wide focused laser spot and collecting the Raman signal in transillumination mode.[116] Spheres of 3nm diameter were resolved by analysis with Lorenz–Mie theory and both modalities successfully registered differences between isolated monocytes and granulocytes.[117] While this combination system shows great promise for single cell applications, it is not well suited to a high throughput cancer screening application.

In our proposed final combined system, elastic scattering will be exploited for its

speed as a pre-screening process, aiming to coarsely segment the tissue. It is envisioned that each of the segmented areas will then be randomly point-sampled using Raman scattering to make the tissue diagnosis, employing the high chemical specificity and proven cancer diagnostic capabilities of Raman spectroscopy.

Both hyperspectral elastic scattering microscopy and inelastic scattering based spectroscopy have been shown to be sensitive to cancerous changes in tissue. This does not, however, mean that they are sensitive to the same biological alterations. It is possible that the segmentation afforded by the HSI will be very different to the Raman maps and the two techniques will not be complementary for this application. In this case, its performance against the gold standard, histopathological evaluation of H&E stained tissue, will be evaluated alone. The HSI-based pre-screening process may have merit as a solo technique for oral cancer tissue section screening as research has shown that unstained HSI is sensitive to the most influential diagnostic features due to scattering changes produced by altered tissue architecture and morphology during neoplastic transformation.[102]

Head and neck, and especially oral, tissue is notoriously diverse due to its complex anatomy.[79] Naturally, oral lesions are also complex and varied making it difficult to discriminate cancerous tissue from healthy due to the large intra-group variations. Even when isolating nuclei, characteristics of which encode more cancer information than other sub-cellular components, cancerous spectra are difficult to distinguish from healthy. Ma *et al* were unable to make this separation in one fifth of their cohort of head and neck (larynx and hypopharynx) cancer samples, yielding specificity ranging from 51 - 82 % dependent on analysis method. This once again highlights the importance of fully exploring the data by using a range of analysis methods. To this end, we will explore both spectral and spatial analysis techniques, performing PCA for dimensionality reduction, k-means clustering for spectral-based segmentation and also texture analysis based on GLCM statistics. These techniques have performed well in previous hyperspectral scattering work and require very little user input, making them ideal for use in an automated clinical system.

1.3.1 Thesis Outline

This thesis will outline the development of a HSI oral cancer tissue segmentation system, for use in combination with Raman spectroscopy, for the detection of oral cancer from unstained, frozen tissue sections. Each of the three Parts details data collected with, a different system which was improved based on the work undertaken in previous Part(s) and with its intended clinical application in mind. Full data collection and correction software will be developed, along with a coregistration workflow to allow

comparison between the three segmentation modes; histopathology (gold standard), HSI and Raman spectroscopy. User input will be minimised in this workflow with the aim of complete automation of the final system and dissemination of associated software for use by others in the field, where possible. The performance of the proposed HSI tissue pre-screening segmentation system will be evaluated with reference to consensus histopathologist diagnosis, and compared with the segmentation achieved by the Raman spectroscopy to assess their compatibility in a combined system. Potential application of the HSI or HSI-Raman system for the screening of bioaccumulation of particulate matter and microplastics will be explored briefly in Part II, followed by overall conclusions and recommended future work.

Part I

Initial HSI System Development and Proof of Principle

As there was no documented system that successfully used visible wavelength HSI on unstained oral tissue sections for oral cancer detection, proof of principle was sought using this preliminary system (Figure 12) at varying stages of its automation.

2 Early Setup and Software Development

The first test system was a transmission mode spectral scanning HSI system, comprised of a tungsten halogen source , a monochromator (CambridgeLifeSciences Ltd. SpectraMASTER) with its "monochromatic" (~ 5 nm spectral resolution) output light coupled via a 50:50 beamsplitter (Semrock) to a microscope (Zeiss Axiovert) with a CCD camera (Hamamatsu ORCA-03G digital CCD) collecting the widefield images of the sample. As the monochromator software was no longer supported, the wavelength was initially altered manually.

Two different samples were inspected with this system to assess the potential of the technique, and to identify necessary alterations to the first, basic system. 505/515 nm FluoSpheres™were dropcast onto a glass microscope slide and dried before being imaged with a x20, 0.5 NA objective (Olympus) in 20 nm steps across the full wavelength range of 350-750 nm. Also, an unstained section of submandibular gland was imaged with x10, 0.25 NA objective (for an increased field of view) at a fixed exposure time of 70 ms, in 20 nm steps across the full wavelength range. The resulting hypercubes were then input into the home-built cluster analysis software that extracts and compares the transmission spectrum for each pixel of the full image, at each wavelength using k-means overclustering and principal component analysis (PCA) as detailed in 3.3.

Fluorescent beads were accurately highlighted by the cluster analysis, however small background contributions were included in the same clusters. This is likely due to the large wavelength step size inhibiting feature separation. Additionally, the centres (x,y) and tops (z) of the beads were members of a different cluster to the edges (x,y) and middles (z), which shows, as expected, a change in focal plane with different illumination wavelengths which requires correction or compensation.

In the unstained tissue hyperstacks, different wavelengths clearly highlight different features. For example, ducts are clearly visible at 400 nm but are very faint by 500 nm as shown in Figure 9. These pixel intensities which change with illumination wave-

length evidence transmission spectral features which could either be caused by photon absorption or scattering events: As the samples are unstained and colourless and the spectral features are broad, the former is most probable.

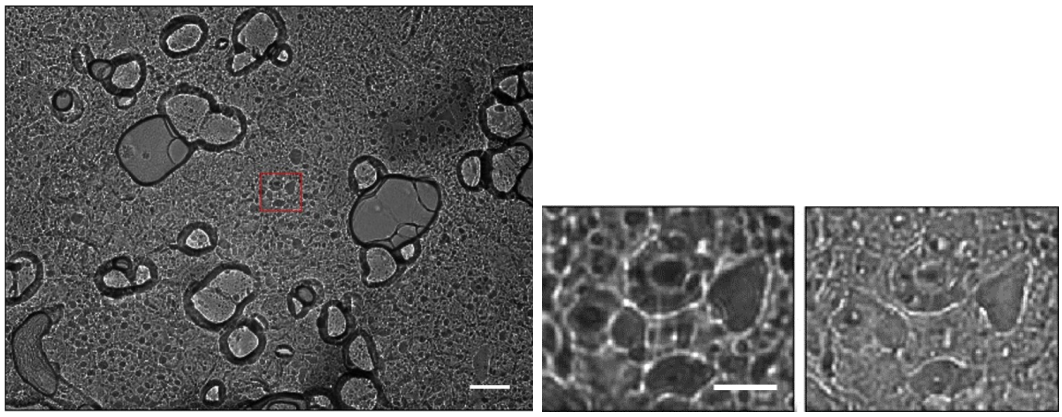

Figure 9: Top: An image of unstained submandibular gland tissue with x10 objective at 400 nm from hyperspectral datacube. Red box highlights a duct and the scale bar represents ~ 60 µm. Bottom: An enlargement of that duct at 400 and 500 nm with scale bar representing 15 µm.

Whilst different features in the tissue sample can be distinguished, decreased wavelength step size may be required to ascertain the source of the spectral features (absorption or scattering) and to produce more precise and distinct spectra to facilitate more informative spectral segmentation. The coarseness of the spectra is evidenced in Figure 10.

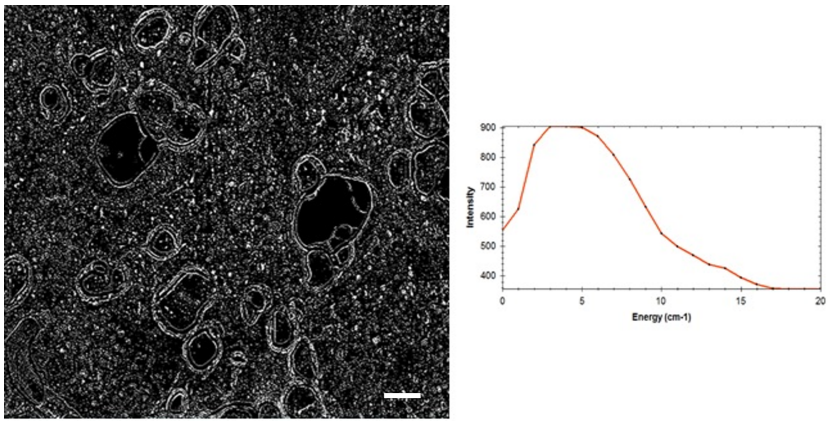

Figure 10: An output from the cluster analysis software calculated from the aforementioned hypercube, collected across is 350 - 750 nm (step size 20 nm) with exposure time of 70ms, showing pixels (in white) contributing to the cluster spectrum on the right. The scale bar represents ~ 60 µm.

Accuracy of the spectra produced were also compromised by source power variation across the wavelength range which highlighted the need for exposure time correction.

In order to ameliorate the focus, segmentation and exposure limitations identified by these preliminary experiments, it was necessary to automate the monochromator wavelength sweeping. An Arduino Uno microprocessor with an Analog Shield (Digilent, National Instruments) was used to provide DACs to generate an output voltage within a ±5V range which, when connected to the monochromator, altered its wavelength output. This made it more practical to utilise smaller wavelength increments and to automate basic focus Z Motor (Prior Scientific) and exposure time correction routines.

Autocorrelation-based autofocus and iterative autoexposure routines performed well but substantially extended the data collection time. A compromise between this improved data quality and the increased data collection run time was therefore made: The autofocus and autoexposure routines were each run once and the optimal focal plane positions and exposure times for each illumination wavelength were saved in a simple text file. For each illumination wavelength during data collection, these values were read in from their respective text files, ascribed to the focal position and exposure time variables and utilised accordingly resulting in significantly faster imaging and was employed for all systems in this work. Fluctuations in illumination and Z-stage shifts or drifts between datasets were therefore not accounted for. To minimise the effect of these small changes between datasets, all systems were made light-tight and employed objective lenses with large focal depths.

These alterations minimised the three main issues discovered during these preliminary experiments but also highlighted the need for a light source spanning a broader wavelength range in addition to a monochromator with software control which operates across this range, leading to the development of the first proof-of-principle system.

3 Preliminary HSI System and Software

3.1 Setup Development

The improved monochromator required more careful installation and alignment and to this end, a completely separate preliminary setup was built on an optical bench, shown in Figure 12. This second system was a transmission mode spectral scanning HSI system comprised of a microscope (Zeiss Axioskop) coupled to a white light source (ThorLabs OSL1-EC Fibre Illuminator, ~350-1000 nm) and broader range, fully automated monochromator (Newport Corporation Oriel Instruments Cornerstone 130, 300-2100 nm with Newport 74027 ruled diffraction grating, 600 l/mm, 700 nm blaze, 450-2000 nm) and camera (Basler ac2040-90um sCMOS). Both the monochromator and camera had drivers provided which were used in the control software, written using VB.NET.

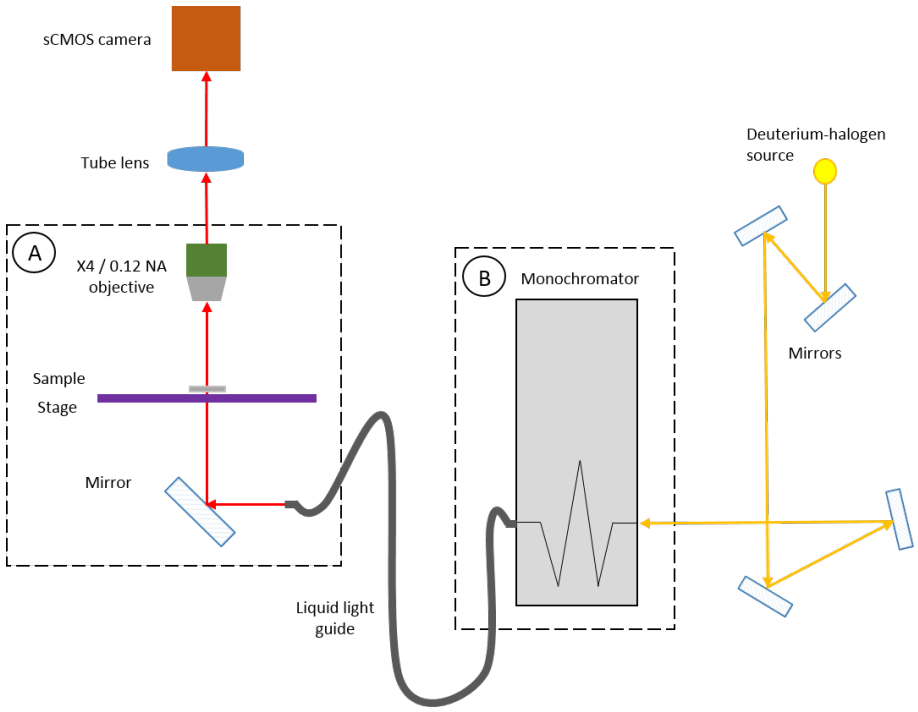

Figure 11: The light path of the preliminary HSI system with labelled areas A and B corresponding with the photographs in Figure 13.

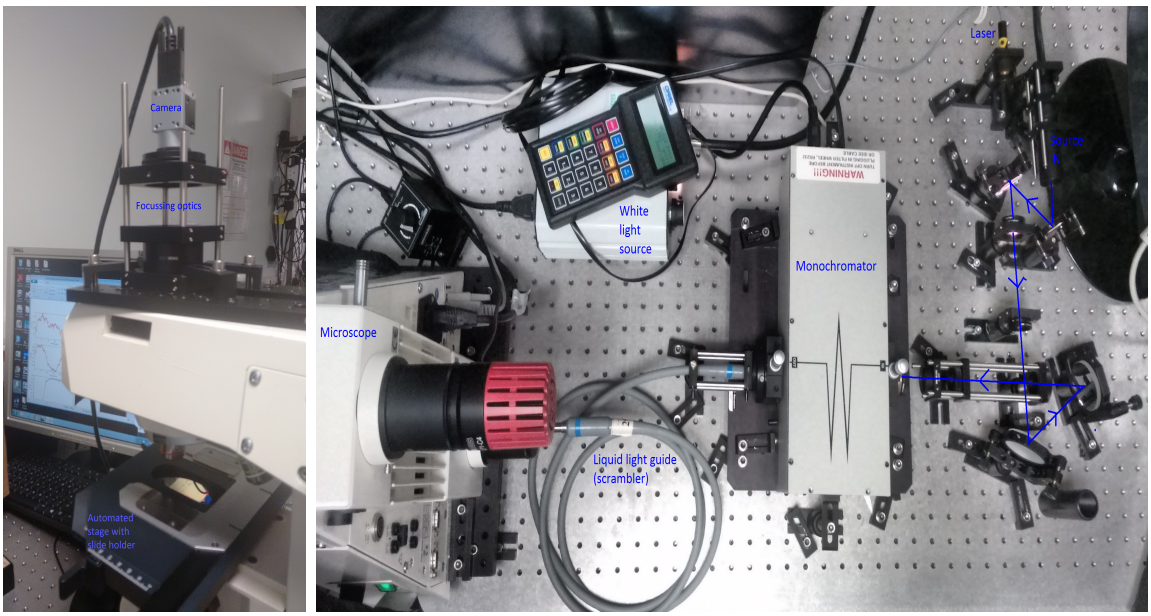

Figure 12: A photograph of the preliminary system from above with light path.

A flipping mirror was used to switch between the alignment laser and white light source to allow alignment to be monitored. The series of alignment optics before the monochromator entrance slit were initially mirrored identically after the monochromator output slit. However, test datasets showed image tracking in x and y across the

CCD with change in wavelength, most likely due to chromatic aberration of a slightly off-centred beam. Without significant manual pre-processing, pixelwise spatial analysis would be impaired as subsequent images in the hypercube would not be directly comparable and the effective spatial resolution of the images would be reduced to the tracking distance in the x and y directions. The optics used to launch the light into the microscope were therefore replaced by a fibre (Thorlabs M37L 550 µm 0.22 NA Step-Index Multimode Fiber Optic Patch Cables: SMA to SMA) and a liquid light guide, before and after the monochromator respectively, to homogenise the light source and a mirror (Thorlabs PFR10-P01) mounted at a 45° angle to direct the light up through the sample. This alteration led to a new illumination path, shown in Figure 13, that circumvented the internal microscope optics. Most lenses are only chromatic aberration-corrected forwavelengths in the visible range, leaving the extremities of our illumination range vulnerable to aberrations and the associated loss of power. This new light path almost entirely consisted of reflective optics which do not rely on refraction and therefore avoidaberration.

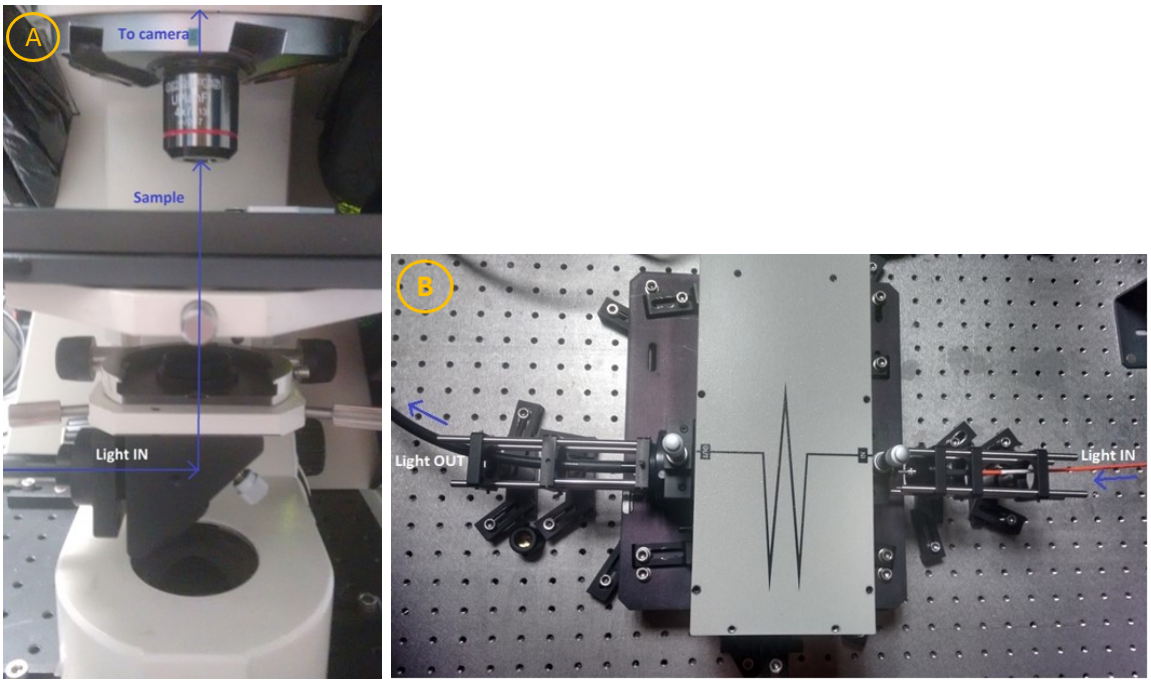

Figure 13: From left to right. The liquid light guide and mirror interface with stage. The simplified fibre-based optics launching light into, and collecting light exiting, the monochromator. This figure corresponds with the sections of the optical path labelled A and B in the optical path Figure 11.

3.2 Data Collection Software

Control of the monochromator, flipping filter, Z-stage and camera was achieved with VB.NET (Microsoft) and developed into one piece of software. VB.NET was chosen

in part due to its automatic graphical user interface generation which makes the software easy to use and train others on which would be an important aspect of clinical translation, preventing the need for an expert microscopist to operate the system on a day-to-day basis, or lengthy training of pathologists to this end. This hardware control software facilitated the collection of a full hypercube with user-specified wavelength inclusion, use of pre-defined focal positions at each wavelength (or autofocus, performed using autocorrelation), use of pre-defined exposure times at each wavelength (or autoexposure), filter flipper control for exclusion of second order light from the monochromator and concurrent dark field stack collection.

Background Correction The illumination being launched close to the sample by the liquid light guide led to difficulties producing a flat field. Test datasets showed that the resulting bright and dark regions impeded accurate segmentation and, to account for this, a more comprehensive background removal protocol was developed and added to the VB.NET control software and workflow.

The correction was based on the standard flat-field correction described by equation 9 but evaluated cumulatively, in sections.

$$C = \frac{m(R - D)}{(F - D)} \tag{9}$$

where C is the corrected image, m is the image-averaged value of (F-D), R is the raw uncorrected image, D is the dark frame and F is the flat field image. The factor (m) is collected and applied automatically as part of the exposure correction routine. The dark frame captures the unilluminated view of the sample, including ambient light contributions and camera "dead pixels" which are assumed to be constant. The flat field images capture the illuminated view of the light path without the sample in place, at each wavelength in the range. They predominantly include vignetting effects and artefacts on, or in between, optical components which can appear different as the illumination wavelength is altered. Flat field images are also afflicted with artefacts isolated in the dark field, which is reflected in Equation 9 and informed the order and frequency of their collection which is detailed below.

The dark frame collection was added to the VB.NET control software: After each raw image (R) the monochromator shutter was closed and another frame was captured (with no illumination) (D) and added to a corresponding dark frame stack which was subtracted from the raw hyperstack using ImageJ at the end of its collection. With the monochromator shutter alternating between open and closed, and with no sample on the stage, a flat field hyperstack (F) and its corresponding dark field were collected across the full wavelength range. These were re-collected for each setup permutation

such as following changes to the light path or objective and the denominator (F-D) was calculated using ImageJ. The manual ImageJ stack subtractions were not integrated into the VB.NET control software as it would have greatly increased the run-time of the software, increasing the likelihood of an incomplete run due to hardware instability. This would be a simple modification to make in a more reliable future system.

Second Order Contamination

Monochromators use diffraction gratings to select a specific wavelength (or waveband) from a broadband light source. The angle of diffraction of each wavelength component of the incoming light can be calculated simply using Equation 10, shown below.

$$m\lambda = dsin\theta \tag{10}$$

where m is an integer describing the diffraction order, λ is the wavelength of the illumination whose diffraction is being considered, d is the line spacing of the grating, θ is the angle between the normal vector of the grating and the intensity maximum of the diffracted ray.

Evidently, for a broad wavelength range, the diffraction angles occupied by subsequent orders may not be unique.

Therefore, when setting the desired output wavelength (λ) of a Czerny-Turner monochromator, higher order radiation (λ/2, λ/3 etc. where the denominator is the order of diffraction) may contribute to the output, along with the first order radiation (λ/1). When the contribution from this higher order radiation becomes non-negligible, output light will not be truly monochromatic and therefore images collected with this illumination will contain additional features highlighted by interactions of that higher order radiation with the sample. This is unsuitable for spectral scanning hyperspectral imaging which relies on distinct, monochromatic feature extraction. Higher order contamination more commonly becomes problematic at the extremities of the wavelength ranges catered to by gratings, due to the increased angles of incidence. In this setup, second order contamination of ~400 nm became noticeable when the desired output wavelength was set to ~800 nm. In order to remove this, a long pass filter (Semrock FF01-776/LP-25 776 nm blocking edge BrightLine® long-pass filter) was inserted into the optical path at the earliest opening in a flipping mount (Thorlabs MFF102/M Motorized Filter Flip Mount with Ø2" Optic Holder). This juncture was after the monochromatic light from the liquid light guide was directed up towards the sample by the mirror. The VB.NET control software was modified to introduce the filter from 785 nm to 950 nm and remove it for all wavelengths lower than 800 nm.

3.3 Preliminary Analysis Software

During the preliminary investigations the source of image contrast was unknown and therefore the optimum analysis technique was also unknown. A previously developed cluster analysis programme that was developed in-house to isolate and map the location of different tissue regions from Raman spectroscopy data was determined to also have some merit in separating regions in the hyperspectral data. This is not unexpected as Raman data and hyperspectral data are similar in many ways; principally, their spectrum-per-pixel format. This software utilised PCA-based noise filtering and k-means clustering to ensure a good degree of similarity within all clusters, followed by a unique overclustering final step which estimates the number of components automatically meaning no prior information about the sample was required.

3.3.1 PCA and K-means with Overclustering Theory

A new technique was created at KCL, with further development by Renishaw, called overclustering which utilises the long established k-means clustering as a subroutine. It works as follows.

1. **PCA Noise Filtering**

   (a) Principle components are curves which show the variance between the spectra. They were calculated by finding the eigenvectors of the covariance matrix for all the spectra at each wavenumber (or wavelength). The eigenvectors themselves are the principle components.

   (b) They were then ranked according to how great an effect they had on the recalculation of the dataset and assigned a weighting accordingly.

   (c) To eliminate noise, only the first few principle components were used to recalculate the spectra in the synthetic (noise-filtered) dataset.

2. **Overclustering**

   (a) The k cluster centroids (mean spectrum of the cluster) were initialised using Monte Carlo. How "close" each spectrum was to this centroid was then calculated in terms of the Pearson Correlation Coefficient starting at k = 2.

   (b) Each spectrum was then assigned to the cluster with the closest centroid, shown by the largest PCC value (r_{max}).

   (c) The spectrum with the smallest value of r_{max} (r_{low}) (the least similar spectrum to the two centroids) was then selected to seed a third centroid.

(d) K-means was then repeated with k = 3, and so on, until r_{low} changed less than a set amount. (This eliminates randomness: k-means has randomness throughout whereas the only random step of this overclustering method is the seeding of the initial centroid. From k = 2, the method of seeding the new clusters is systematic which ensures that the same set of data consistently produces the same clustering results.)

3. **Optimal Cluster Number Selection**: This step automated selection of the optimum number of clusters, k, ideally corresponding to all of the separate tissue constituents. If there are k unique centroids which make up an optimum set, overclustering will find this set and a number of additional clusters. As the optimum set must cover the full range of the tissue constituents, any additional clusters must be linear combinations of those in the optimum set. This simple fact was used to estimate k within an error of ±2 clusters using the following method.

   (a) For each centroid, all of the other centroids were used as a basis to calculate the least squares fit.

   (b) The sum of squares error (SSE) of this fit was then evaluated for each centroid.

   (c) The centroid with the lowest SSE was ranked lowest and removed from the set.

   (d) The first three steps were repeated until only two centroids remained in the set.

   (e) Simultaneously, a graph of SSE against number of centroids remaining in the set was constructed. This graph follows a rapidly decreasing trend which sharply levels off. Centroids below this levelling-off value will be unique.

   (f) The point at which the plot flattens must was then identified manually: This process is somewhat subjective as it very rarely occurs at one point, but usually over two or three.

4. **Cluster Assignation**

   (a) Each of the optimum centroids (as found above) then underwent a least squares fit with the others in the set as a basis.

   (b) The least squares coefficients were normalised to one and then used to calculate the probability that a particular spectrum is a member of a cluster given by a specific centroid.

(c) The highest cluster probability score was then used to assign spectra to a cluster.

(d) Each cluster channel was then represented visually by plotting the pixels with spectra assigned to that cluster in white, on a black background.

(e) Colour maps could then be generated by randomly assigning a hue to each cluster and plotting the position of all clusters in their respective colours on one image.

Output

The segmentation results from the cluster analysis were output in a tabular format shown in Figure 14, including the spectrum of each cluster and the distribution of those clusters for each sample. Each cluster spectrum is shown as a red line connecting black data points with axes of (pixel) intensity and frame number, the latter of which ranges from 0 - 300, representing each illumination wavelength of the hyperstack; 450 - 750 nm. Showing the cluster distribution in each sample side-by-side allowed comparison between the distribution of individual spectral components in different samples, assisting identification of the component. Additionally it was possible to output a colour-coded map condensing all of these segmentation information into one image per sample, and a corresponding coloured pie chart showing the the relative areas of each sample described by each cluster. These representations of the data were useful, however in both instances the colours were randomly generated, which occasionally jeopardised the clarity of the colour maps.

		Sample 1	Sample 2	...
Channel 1	Spectrum 1	(a)	(b)	
Channel 2	Spectrum 2	(c)	...	
Channel 3	Spectrum 3	⋮		

Figure 14: A table illustrating the format of the results output by the home built overclustering software used for the analysis of HSI and Raman data collected using the early and preliminary HSI systems detailed in Part Is. The spectrum is a plot of intensity with wavelength (counting from 0 at the start wavelength) which is, practically, a plot of the grayscale value at each pixel in each slice of the hypercube. The channel row shows where the corresponding spectrum is present in each sample loaded into the analysis software simultaneously. The entries (a), (b), (c) etc. show a binary map of the sample with white pixels indicating the location of that spectrum (row) in that sample (column) scan area. For example, spectrum distribution map (a) shows where spectrum 1 can be found in sample 1, (b) shows where spectrum 1 can be found in sample 2, (c) shows where spectrum 2 can be found in sample 1 etc.

4 Materials and Methods

4.1 Sample Selections and Sourcing

A variety of samples, outlined below, were selected to investigate different aspects of the system's capabilities.

Samples	Rationale
Fluorescent Dyes	Well characterised samples to test system spectral accuracy.
Blood	A well characterised biological sample to ascertain the ability of the system and cluster software to identify changes in concentration.
Stained Oral Tissue	To test the performance of the HSI system on an enhanced (visible wavelength stained) sample of the target tissue.
Whole Mouth Saliva	A transparent but complex biological sample to ascertain whether the system could highlight any features invisible to the naked eye.
Unstained Oral Tissue	Preliminary proof of principle and to assist software development and to test the viability of visible wavelength illumination HSI of unstained tissue.

Table 1: A table outlining the rationale behind the proof of principle samples selected to investigate the capabilities of the preliminary hyperspectral system (Figure 12).

Reference spectra of fluorescein, rhodamine and blood are included below for verification of the spectral accuracy of the system. The centrepoint of the rhodamine and fluorescein dye absorption peaks (shown in Figure 15) lie at 560 nm and 488 nm, respectively. [118]

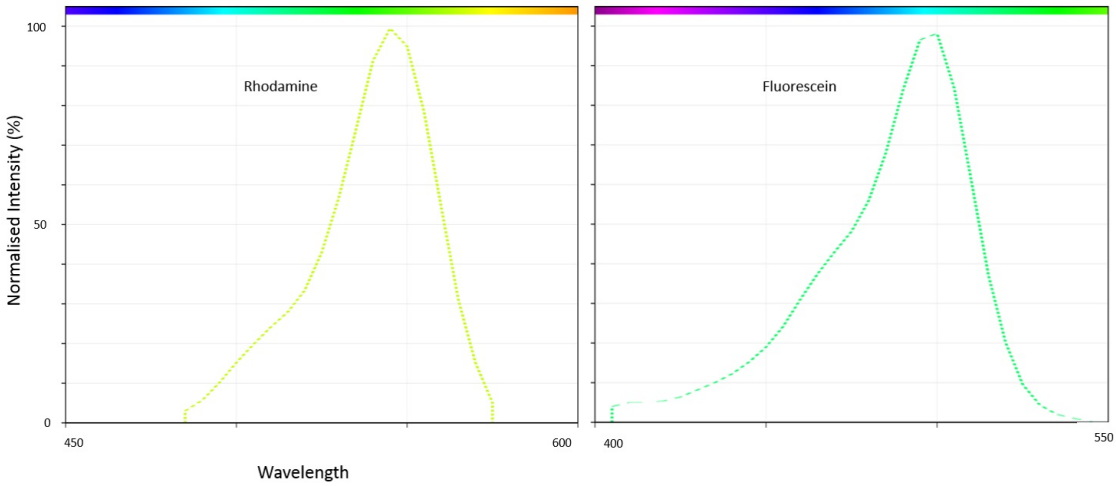

Figure 15: The absorption spectrum of rhodamine dye (left) and fluorescein dye (right) made with AAT Bioquest.

The absorption spectra of oxygenated and deoxygenated haemoglobin shown below in Figure 16 was created from data from Prahl *et al.*[119]. The peaks would decrease in intensity on dilution with PBS.

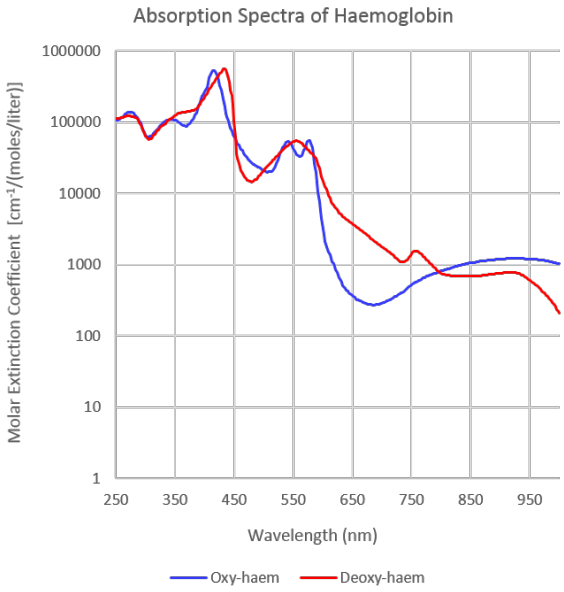

Figure 16: The absorption spectra of oxygenated and deoxygenated haemoglobin.

As the blood was collected using the finger prick method, it is capillary blood which is a combination of arterial and venous blood which are oxygenated and deoxygenated, respectively.

Patient tissue samples and data were provided by Guy's & St Thomas' Head & Neck Biobank – part of the KHP Cancer Biobank, which is supported by the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award and Guy's & St Thomas' NHS Foundation Trust. The biobank has ethical approval from the East of England - Cambridge East Research Ethics Committee (18/EE/0025).

This study utilised three healthy and three malignant stained, three healthy and three malignant unstained, oral tissue samples and which were randomly selected from the sample set of Dr. Manickavasagam to ensure representation of a range of anatomical sites and malignancies.[120] Imaging was undertaken "semi-blind" in that I was aware of the diagnosis for each tissue section (in order to select appropriate samples) but not what features on the slide led to each diagnosis.

Staining	Health	Sample No.	Description
Stained	Healthy	8545	Buccal mucosa
		3689	Submandibular gland
		780	Parotid
	Malignant	7772	Parotid acinic cell carcinoma
		4025	Squamous cell carcinoma, tongue
		5178	Carcinoma, ex pleomorphic adenoma
Unstained	Healthy	232	Submandibular gland
		4984	Parotid
		6111	Floor of mouth
	Malignant	8883	Precancerous adenoma, parotid
		8460	Squamous cell carcinoma, oral
		9120	Adenoid cystic carcinoma, parotid

Table 2: A table including a description of of each tissue sample, its sample number and whether it is stained or unstained.

4.2 Sample Preparation

4.2.1 Fluids

The fluids rhodamine, fluorescein, blood (3 different dilutions with PBS) and whole mouth saliva were all prepared similarly by dropcasting directly onto a glass microscope slide and leaving to air dry. When the samples dried, they did so in a coffee ring formation; an effect driven by capillary action in particle-containing solutions. The samples were therefore dried before use to avoid their drying over the course of the imaging process, changing over time as the "coffee rings" form and rendering images at different wavelengths incomparable to one another. The volumes dropcast in each case were 1 ml, 1 ml, 5 µl and 5 µl, respectively. The blood was diluted with phosphate buffer solution (PBS), in order to maintain the structure of the red blood cells, in blood to PBS ratios of 1:1, 1:2 and 1:3.

4.2.2 Tissue

Freezing of Fresh Tissue

1. Fresh tissue was harvested from specimen.
2. This fresh tissue was then placed on a piece of cork with OCT.
3. Frozen in iso-pentane cooled with liquid nitrogen.
4. The tissue was then quenched quickly in liquid nitrogen to ensure complete freezing with no ice crystal artefacts.
5. The frozen tissue sample was then placed into a labelled cryovial and stored in -70°C freezer.

Frozen Tissue Sectioning

The frozen sample was placed on a cryostat chuck and adhered with OCT, which acts as a glue/support medium, by freezing in liquid nitrogen. Care was taken not to thaw the tissue during this process.

From each sample, two sections were cut on Leica CM1850 cryostat:

1. H&E
   (a) A 4 µm thick section was cut and picked up onto a glass microscope slide (Surgipath microscope slide: 3808122GCE).
   (b) The slide with section was placed in coplin jar with frozen section fixative (Formaldehyde (40%): Genta Medical EC 200-001-8 EC Label UN2209, Acetic acid 99-100% (glacial): VWR 20103.295, Industrial Methylated Spirits (IMS): Genta Medical EC 200-578-6 EC Label UN1170) for 2 minutes.
   (c) The section was then stained by H&E (see staining protocol 4.2.2).
   (d) A coverslip (Menzel-Glaser, 22x26 cover glass, 6776311) was then mounted (Histolab pertex 00801-EX) on the slide, covering the tissue.
2. Laser section
   (a) A 10 µm thick section was cut and carefully placed onto a calcium fluoride slide (Crystran CaF_2 20 mm diameter, 0.5 mm thickness, polished Raman grade).
   (b) A second CaF2 slide was gently placed on top of the first, with the tissue section between the two. They were adhered with a thin layer of clear nail varnish ensuring this did not get drawn into the section by capillary action.

(c) The glass slide on top was labelled.

(d) The slides were then stored at -70 to -20 °C.

Staining

H&E slides (slide 1 from sectioning protocol 1) were stained on the linistainer (Shandon Linistain) using the following sequence and time periods:

1. Industrial methylated spirits (IMS) (Genta Medical EC 200-578-6 EC Label UN1170) (20 seconds)
2. Tap water (20 seconds)
3. Harris Haematoxylin, non-acidified (Shanon - Harris Hematoxylin, Non-acidified Thermo Scientific 6765001) (40 seconds)
4. Tap water (20 seconds)
5. Acid alcohol comprised of 0.25% hydrochloric acid (Fisher scientific Hydrochloric Acid S.G UN1789) and IMS (Genta Medical IMS EC 200-578-6 EC Label UN1170) (20 seconds)
6. Tap water (20 seconds)
7. Saturated lithium carbonate solution for bluing (Fisher Chemical Lithium carbonate L/2100/50, in deionised water) (20 seconds)
8. Tap water (20 seconds)
9. Eosin Y 1% (Cell Path Eosin Y Stain 1% (Aqueous) RBC - 0100-00A) (20 seconds)
10. IMS (2 x 20 seconds)
11. Xylene (Genta Medical Xylene EC 215-535-7 EC Label UN1307) (2 x 20 seconds) .

4.3 Data Collection and Processing

All of the HSI datasets were collected as a single field of view (no mosaicing) using a x10, 0.25 NA objective and across a wavelength range of 450-750 nm with spectral step size of 1 nm. In cases where the full sample could not be imaged in one field of view, the area with the most valuable information (assessed qualitatively, by eye) was imaged. The process of how the HSI data collection software was setup up and run is detailed below.

4.3.1 Hyperspectral Imaging

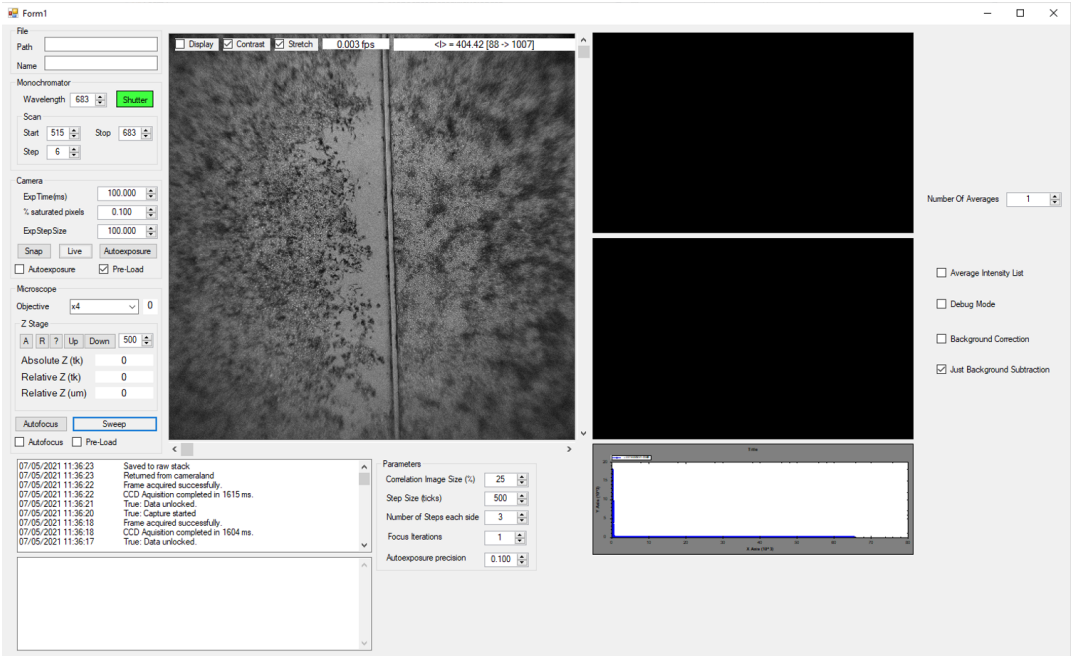

Figure 17: Graphical user interface for the home-built hyperspectral imaging data collection software.

1. The home-built software was launched, showing the GUI in Figure 17.

2. The sample was placed onto the microscope stage.

3. The desired objective was rotated into place.

4. The same objective was selected from the drop down objective menu on the software (this ensured that the correct focal plane and exposure time lists were utilised).

5. The "Live" button was pressed to allow video rate display of the sample for manual focussing.

6. If necessary (for example if focusing at wavelengths which require a high exposure time and for which manual focusing is therefore difficult), the size and number of z-steps could be set at this stage and the automated autofocus routine then instigated by pressing the "Autofocus" button.

7. The following checkboxes were ticked by default: Display, Contrast, Background subtraction, Pre-load Exposure, Pre-load Focus but could be altered at this stage if necessary.

8. The minimum and maximum wavelength values of the desired range and the wavelength step size were then entered.

9. The “Sweep" button was pressed to commence the collection of the dataset.

5 Results and Discussion

The calculation process and output format of the segmentation results is detailed in 3.3.

5.1 Fluorescent Dyes

Approximately one quadrant of both the rhodamine and fluorescein dye droplets were imaged according to the standard HSI data collection routine (4.3.1) and then background-corrected (3.2). Both of the resulting hypercubes were input into the cluster analysis software simultaneously to allow mutual clusters to be identified and to aid comparison. As anticipated, the only mutual clusters present in these datasets are self-evidently background or artefactual as opposed to components within the droplet quadrant itself.

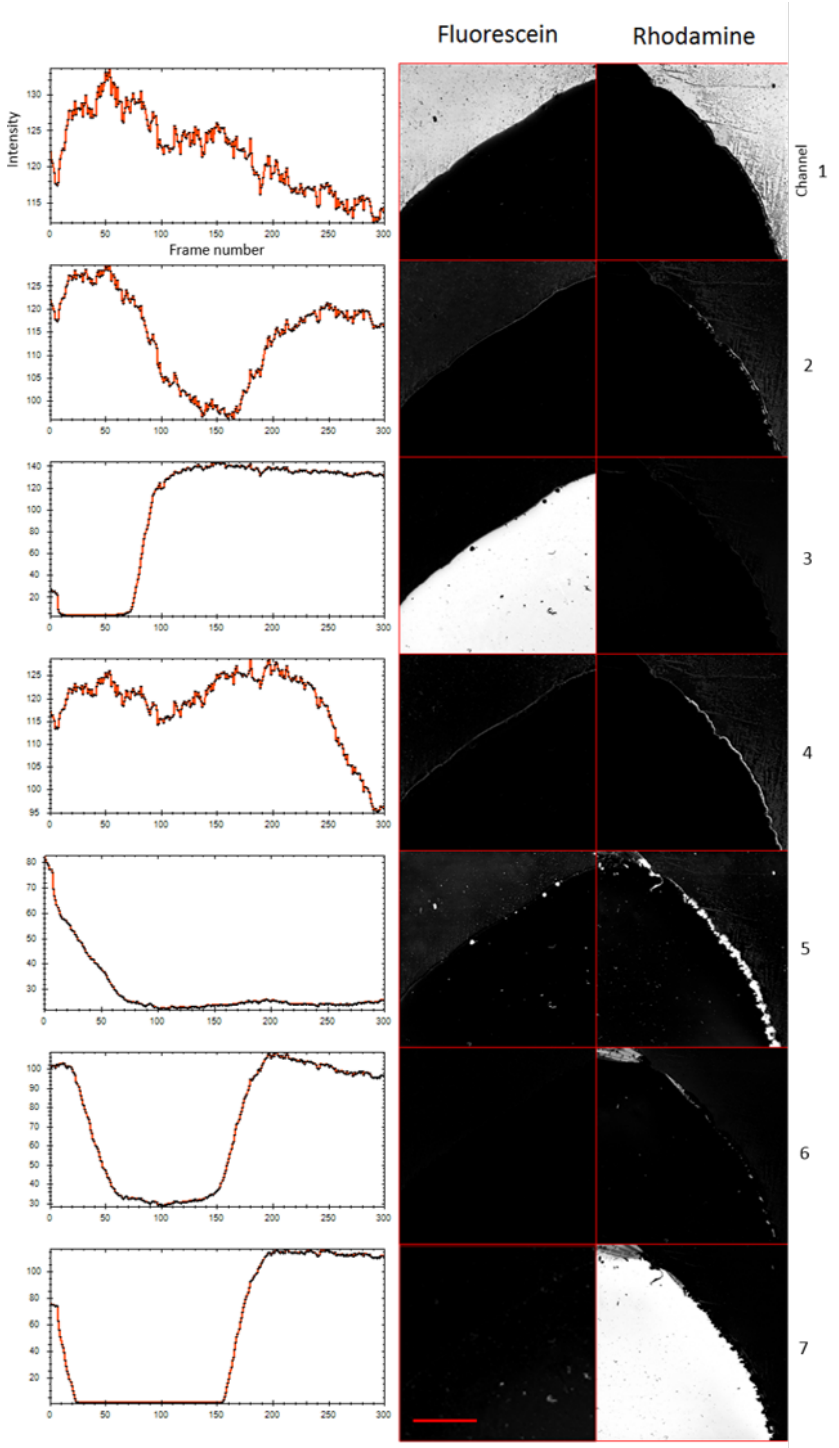

Figure 18: K-means cluster maps of rhodamine and fluorescein. Red scale bar represents 1000 µm.

Referring to Figure 18, the white quadrant in the cluster location map of the third channel (from the top) and the seventh channel evidence that these channels represent the fluorescein and rhodamine drop-quadrants, respectively. This is supported by the first channel which, for both samples, highlights background contributions and which

covers the inverse area to the aforementioned droplet clusters for each dye. To verify this, the spectra of these dye clusters were compared to the reference dye absorption spectra (Figure 15). The cluster analysis software truncated the intensity data and could not be modified. As a result, it was necessary to plot and extrapolate the two linear sections of the spectra from clusters 3 and 7 to find their intersection points, as shown in Figure 19. This provided an estimate of the spectral position of the absorption trough.

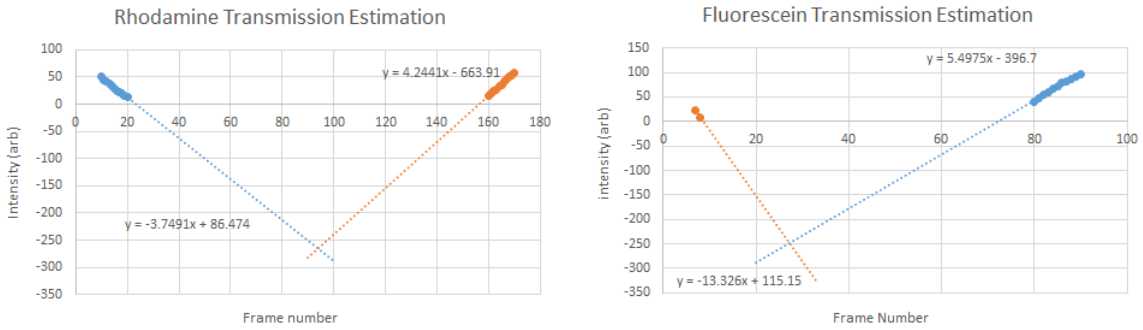

Figure 19: Graph estimating the trough turning points of the transmission spectra from cluster 3 (fluorescein) and 7 (rhodamine). The frame number along the x-axis represent the frame of the hypercube, each of which is illuminated with a different wavelength starting at 450 nm (frame 0) and increasing in 1 nm steps with each frame.

The trough of the rhodamine transmission spectrum from cluster 7 centres around 544 nm which is within one standard deviation of the spectral resolution of the rhodamine absorption peak value (560 nm) of the reference spectrum (Figure 15). Similarly the fluorescein transmission spectrum trough from cluster 3 centres about 477 nm, in good agreement with the absorption peak value (488 nm) of the reference spectrum (Figure 15). Evidently, the spectral accuracy of the system is sufficient and the transmission spectra were correctly formulated and then successfully interpreted by the cluster analysis software.

5.2 Blood

Three 5 µl droplets of blood serially diluted with PBS were imaged in their entirety according to the standard HSI data collection routine (4.3.1) and then background-corrected (3.2). The three resulting hypercubes were input into the cluster analysis software simultaneously to allow mutual clusters to be identified and to aid comparison.

Of the five clusters, three of them (2, 3 and 5) appeared to highlight blood, with each being a different region of the "coffee ring" formation caused by the surface tension variation during the drying of the blood droplets. It was hypothesised that the difference in blood concentration between the drops would be reflected in the cluster analysis

outputs. To verify this, a decreasing percentage coverage with decreasing blood concentration was demonstrated using basic quantitative analysis on ImageJ. The cluster distribution maps were loaded into the ImageJ software. The percentage areas of the field of view covered by each channel were then calculated and plotted for each cluster, for each blood dilution (Figure 20).

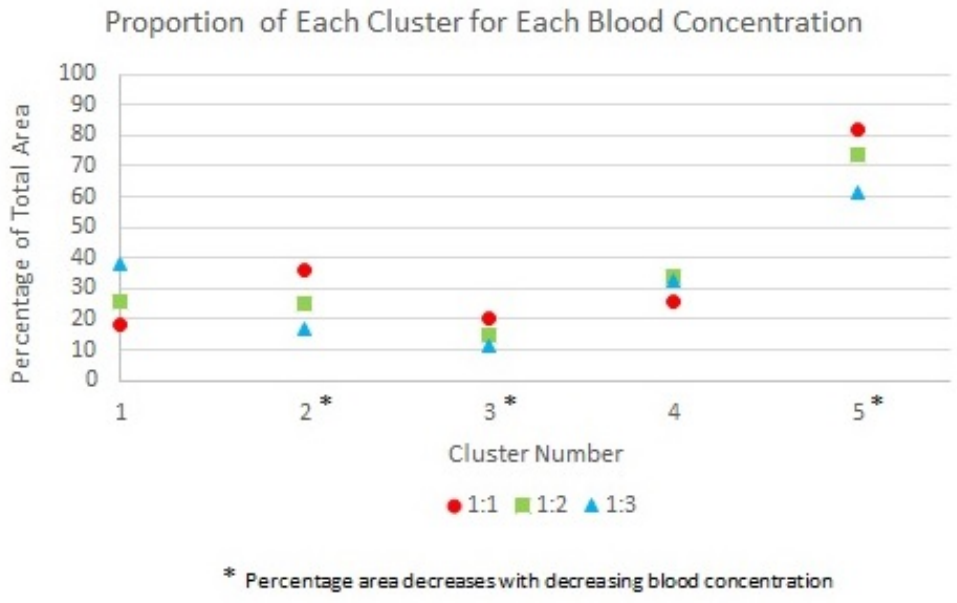

Figure 20: A plot of the percentage of the total image area covered by each of the five clusters for each of the three blood concentrations.

There are three cluster distributions (2, 3 and 5) for which the percentage area decreases with decreasing blood concentration identified by red, green and blue point markers in order from top to bottom. These are the same three clusters that were independently identified as representing blood, showing that the system is capable of detecting change in concentration.

This may be a useful feature of the system as the concentration of various cofactors can be increased in cancerous regions and could therefore be used to help locate malignant regions.[121] Whilst no relevant cofactors are as strongly pigmented as haemoglobin, this result shows that the system has potential to detect and localise changes in absorber concentration.

5.3 Stained Oral Tissue

One central field of view of each of the three healthy and three malignant oral tissue sections stained with H&E were imaged according to the standard HSI data collection routine (4.3.1) and then background-corrected (3.2). After collection, the six resulting

hyperstacks were scaled down to 1022x1022 pixels in ImageJ using bilinear interpolation to enable simultaneous cluster analysis for easy identification of mutual clusters and to aid comparison. After PCA denoising and k-means overclustering, the location map of each cluster for each of the six samples (Figure 21) were output separately in the tabular format described in 3.3. The colour maps show the distribution of all clusters across each sample. The colours are randomly assigned to each cluster and do not encode a known substance however, by analysing all six samples simultaneously, the colour assignation for each cluster is preserved across the samples.

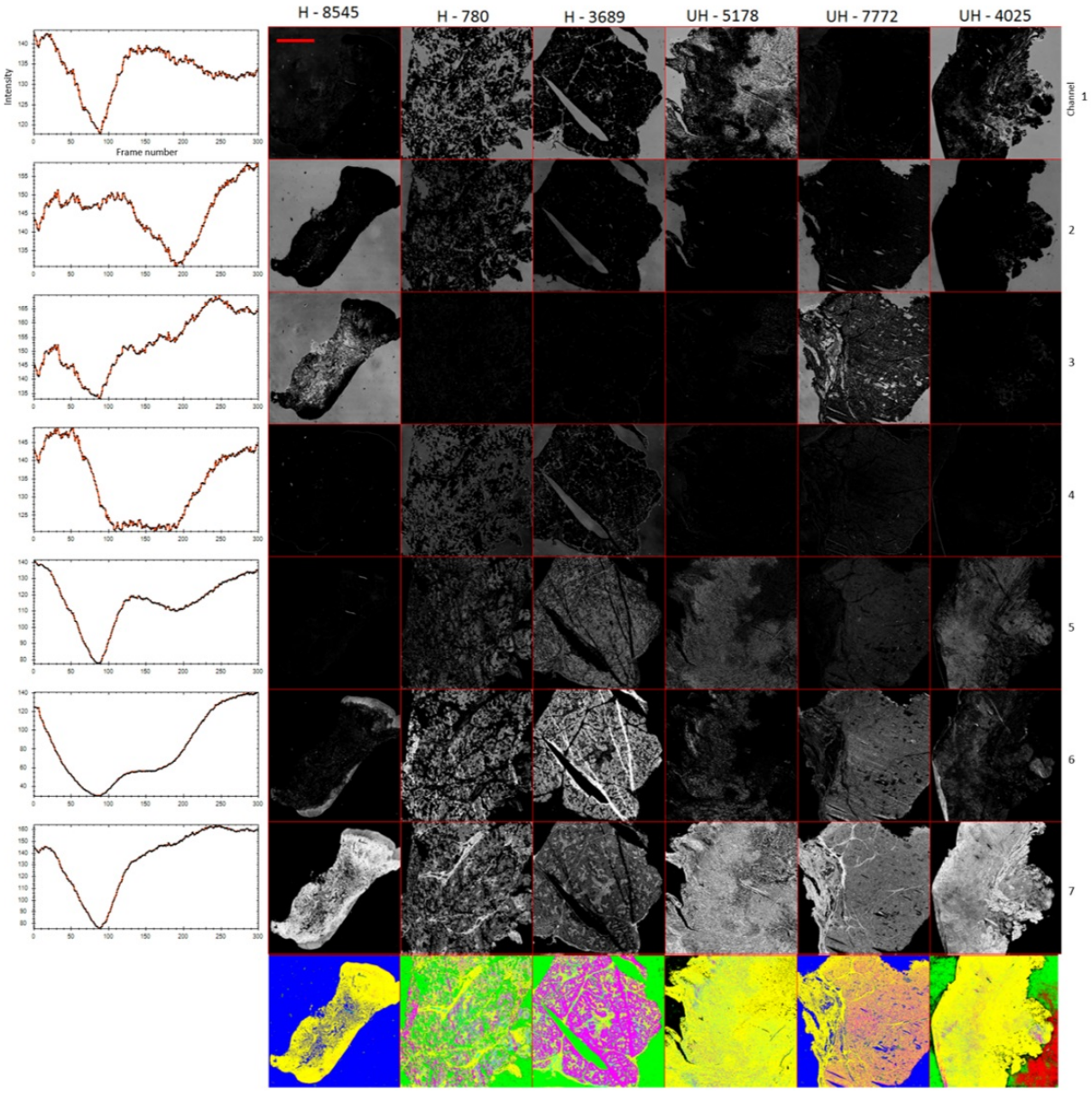

Figure 21: K-means cluster maps of stained oral tissue samples. Red scale bar represents 1000 µm.

The minimum requirement for a pre-screening process would be tissue and non-tissue

separation as this alone would improve the efficiency and speed of Raman-based diagnostics by allowing background to be automatically excluded from the time-consuming Raman data collection process. Tissue and non-tissue regions are accurately separated for all six stained samples. Additionally some features within tissue boundaries are highlighted in all samples, particularly in clusters 6 and 7.

Fibrous tissue is consistently selected across all samples, shown in yellow in the colour maps, and its distribution can be linked to health status. For example, in healthy tissue section 3689, the fibrous runways which take the blood vessels and ducts into and out of the gland architecture can be seen throughout the tissue in cross section. Playing a very different role, the fibrous region on the left of the parotid acinic cell carcinoma sample 7772 is a reaction to the tumour and can be seen in Figure 22 attempting to support the tissue around the tumour as it invades.

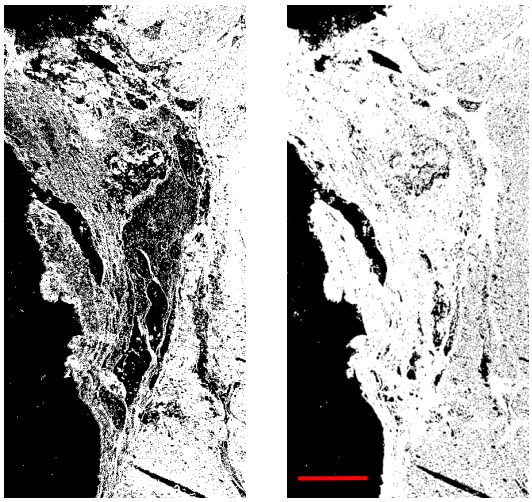

Figure 22: Potential fibrous tissue in sample 7772 shown lacking in the binary map of the distribution of cluster 6 (left) and present in 7 (right). Red scale bar represents 500 µm.

While the tumour is seen predominantly in pink in the colour map of this sample it is important to remember that, on stained tissue, different clusters will group features which are stained similarly, though they may not necessarily be the same tissue component as scattering spectral features will be overwhelmed by strong absorption. Additionally, cancer is made from cells and therefore stains similarly to healthy cells, though with differences in proportions, such as enlarged purple stained nuclei. This means that the distribution and density of the clusters will be more useful in detecting malignancy in stained sections. While the other two malignant samples appear saturated with yellow, the previously identified fibrous component, its distribution aligns with the specific pathology associated with the malignancies being presented. For example, the widespread fibrous architecture in sample 5178, a carcinoma ex pleomor-

phic adenoma, occurs as a response to this tumour which attempts to make epithelial structures everywhere: Healthy epithelial layers such as skin are ordinarily anchored to mesothelium by fibrous tissue and our body responds to this widespread epithelial tumour by trying to producing these conjoining fibres, enacting a familiar response to unfamiliar epithelium. Local fibrotic reactions across the tissue section aso explain the distribution of the yellow cluster in tongue squamous cell carcinoma sample 4025. The body attempts to create a fibrous barrier between the tumour and healthy tissue but, as it invades, it is constantly broken down and remodelled, leaving widespread fibrous residue and poorly delineated tendrils of cancer in the strip of mixed clusters down the middle of the field of view.

The distribution of cluster 2 appears to represent a large component of the background for each sample with the distribution of clusters 1, 3 and 4 showing additional background features for some samples. In an ideal dataset, the background would be entirely described by one cluster. The fact that this is not the case suggests that there is fluctuation or unevenness in the sample illumination and that the background correction is not completely successful in compensating for these. As the dark field images are taken immediately before each corresponding tissue hyperspectral image, this is less likely to be the source of the inconsistencies as the fluctuations would have to occur on a very short timescale. One potential method of minimising background fluctuations in corrected images would be to collect the flat field stack in tandem with the raw and dark field images: Unfortunately this is not possible without a motorised stage, which would allow the sample to be moved out of the light path.

5.4 Saliva

Saliva is a complex biological sample with many components including cells, proteins, minerals, buffers and electrolytes.[122] As the sample is also transparent to the human eye, it is ideal for testing the potential of the system for imaging unstained, colourless samples.

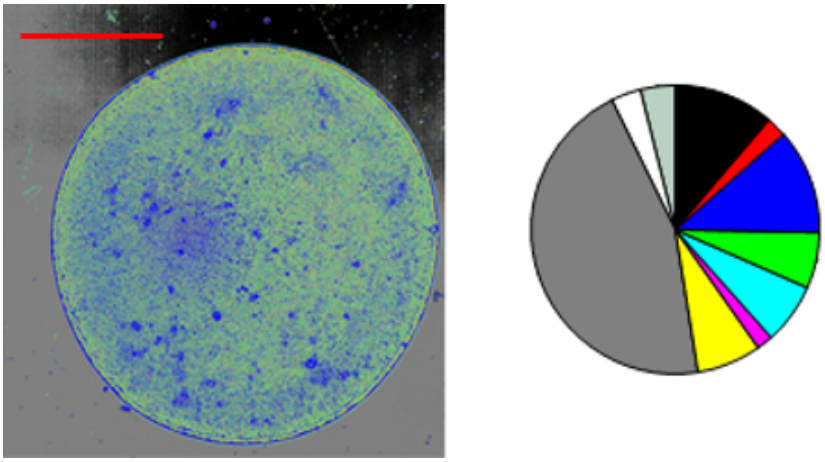

Figure 23: The final map of the clusters found in the hyperstack of the saliva droplet sample and a pie chart showing the relative abundance of each cluster in the sample. Red scale bar represents 1000 µm.

The pie chart in Figure 23 indicates that ten different components were detected in the sample showing that a great deal of extra information can be provided by imaging the sample using HSI rather than the naked eye or conventional widefield microscopy, which are the two most commonly used diagnostic methods employed in histopathological diagnosis from tissue samples (although they would ordinarily be carried out on stained tissue). Vibrational spectroscopy would need to be employed in order to identify the chemical composition of each component but being able to separate and locate individual components beforehand would aid in developing a more time-effective diagnostic data collection process.

5.5 Unstained Oral Tissue

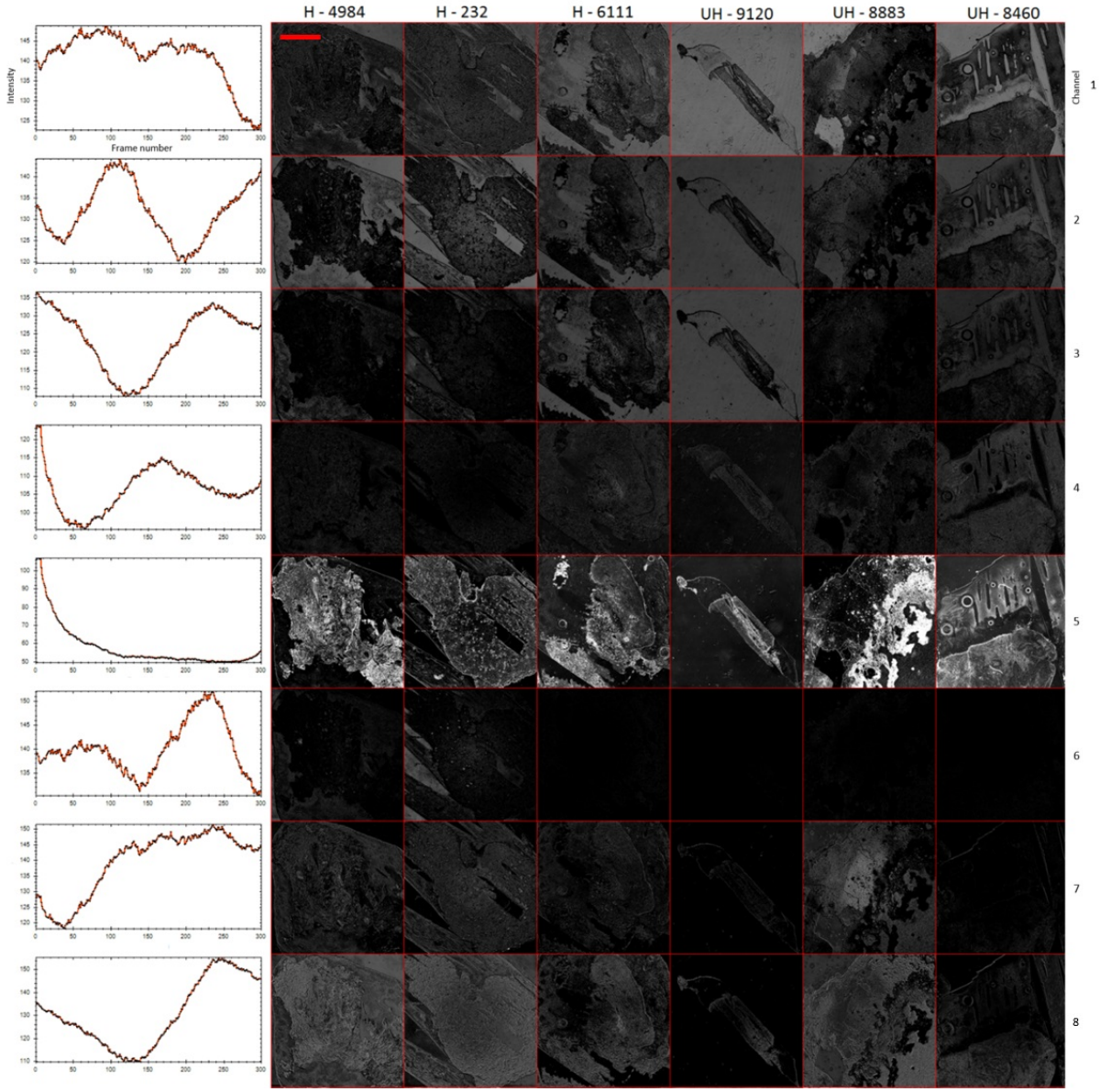

Figure 24: K-means cluster maps of unstained oral tissue samples.

The first three clusters shown in Figure 24 contain background components, with the distribution of channel 3 showing the best isolation of background from tissue. The distribution of cluster 5 isolates tissue features with negligible contribution from background, and its spectrum resembles Rayleigh scattering with a strong wavelength dependence which favours shorter wavelengths. This channel, which appears to isolate scattering signal from absorption effects, contains the best separation of tissue features from tissue. Whilst scattering is segmented out in this dataset, there is no similar channel in the stained dataset. This is likely due to the absence of dominating absorption features in this unstained data.

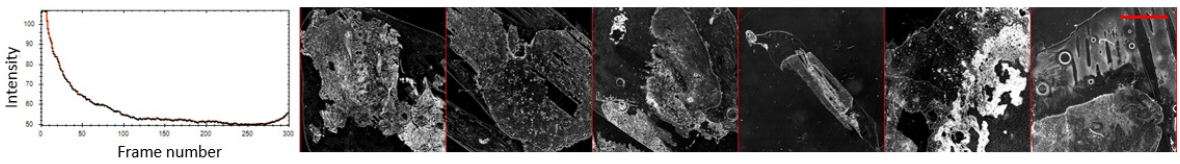

Figure 25: Channel 5 of the cluster analysis output for the unstained tissue, with contrast hypothesised to be provided by Rayleigh scattering. Tissue sample numbers from left to right: 4984, 232, 6111, 9120, 8883, 8460.

Features much smaller than the wavelength of illumination will produce Rayleigh scattering. The distribution of this cluster (shown in in Figure 25) correlates with fibrous structures: Collagen fibrils are ~0.3 µm in diameter with periodic striations of ~70 nm, and therefore chiefly produce Rayleigh scattering of visible light, though Mie scattering of collagen fibres >2 µm in diameter becomes significant in the infra-red wavelength range.[123] Conversely to FTIR where scattering is a hindrance to analysis, this setup and analysis software is clearly capable of separating scattering completely, rendering it advantageous. This scattering could be purposely made more prominent if the setup was altered to utilise longer illumination wavelengths.[124] This cluster is likely to contain information from other similarly sized tissue features, such as lysosomes and vesicles, along with fibres though the fibrous architecture has the greatest potential for diagnostic relevance and is likely to be visible even if future systems utilise a lower magnification for faster slide screening.

As in the stained tissue segmentation results discussed previously, the distribution of this fibrous component demonstrates the expected pathology of the samples. In sample 232 of healthy submandibular gland, we see cross sections of ducts (like in healthy stained sample 3689) and similarly, we may see fibres tracking a nerve in the adenoid cystic carcinoma parotid sample 9120. While this sample is malignant, it has a deceptively benign histologic appearance with a lower density of cancer cells due to cyst-like acellular cavities. Sample 8883 is the pre-malignant version of the cancer of the carcinoma ex pleomorphic adenoma stained sample we discussed previously and, while it is currently benign, this fibrous Rayleigh scattering channel appears to show architectural changes and variation of structures within the tumour. An unrepresentative field of view was selected from sample 8460, an oral squamous cell carcinoma, with the top half showing features within the OCT (optimal cutting temperature) compound rather than the tissue itself: When the OCT contains bubbles or cracks like this, it can be difficult to identify the tissue edge. However, at the bottom of the field of view, the tumour appears to be creeping around the tissue in the middle, which appears brighter due to its innate fibrous support which is enhanced as the body responds to the tumour invasion.

Scattering aside, the distribution of cluster 8 highlights tissue somewhat successfully, as do channels 7 and 4 to lesser extents. Where this tissue segmentation fails often coincides with areas of fixative, which can be observed by comparison of the sample 4984 cluster distribution maps for channels 5 (Figure 24) and 8 (Figure 26). This issue occurs inconsistently between samples in spite of the same fixative being used in each case. One potential explanation for this variation is ageing or degradation of the fixatives which could alter their transmission spectra and which could occur on a sample-by-sample basis due to slight differences in age, storage conditions and number of freeze-thaw cycles.

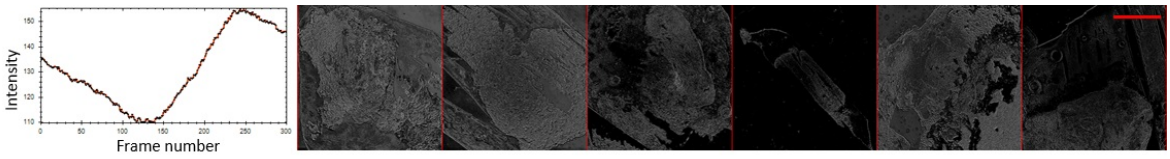

Figure 26: Channel 8 of the cluster analysis output for the unstained tissue, exhibiting successful tissue and non-tissue separation. Tissue sample numbers from left to right: 4984, 232, 6111, 9120, 8883, 8460. Red scale bar represents 1000 µm.

Cluster 2 selects some potentially interesting components such as the central regions of sample 4984 and 232 which correlate with healthy cellular regions. Cluster 8 (Figure 26) also appears to highlight cellular regions, being distributed more densely in areas of increased cellular density such as in the malignant cancer regions in samples 9120 and 8460 and the denser glandular healthy samples 4984 and 232. The pleomorphic adenoma sample, 8883, is named to reflect the variety of structures it can comprise (including keratin, glandular tissue and cellular regions) and its characteristically inhomegenous cellular density is evident in cluster 8. This could suggest that this cluster results from scattering which is affected by cell packing densities and the resulting changes in nuclear density, though the spectrum of this cluster appears to contain features other than the flat or decreasing, oscillatory trend we would expect to see from Mie scattering of small (healthy) and enlarged cancer cell nuclei.

Unstained tissue is essentially colourless and transparent and therefore their transmission spectra do not contain clear absorption features. The majority of contrast in hyperspectral imaging of colourless samples comes from light scattered by boundaries between regions of different refractive indices throughout the tissue. Due to the absence of true background in most images due to the presence of OCT and the small field of view relative to the sample sizes, reliable calculation of the Weber contrast[6] was not possible for this dataset. However, the small dynamic range and relatively even dis-

[6]Weber contrast describes the luminance of the subject being imaged, relative to the image background.

tribution of different pixel intensities within this dynamic range imply that the Weber contrast of the unstained images (or spectra) is low.Therefore the cluster distribution maps are mostly low specificity, with less clearly defined differences between channels than in the stained tissue. Efforts to increase the Weber contrast could improve the segmentation by the cluster analysis. The low spectral resolution of the system was a likely contributing factor to the lack of specificity and resulting uncertainty regarding the clinical relevance of segmentation achieved. The spectral resolution of the system was limited by the small illumination beam width on the monochromator grating (though 1 nm wavelength steps were employed in the hypercube collection for these exploratory datasets) which led to fine spectral features, which may have been valuable in the segmentation process, being under-sampled. Diagnostic classifiers must be identified before the optimum spectral resolution can be identified. Until that point, adequate sampling of all potentially relevant spectral features must be ensured and, to that end, the highest possible spectral resolution should be employed.

This hyperspectral system and segmentation software has produced segmentation with better separation of sample from background and more useful biological information than a widefield image taken with a single illumination wavelength; this is due mostly to the Rayleigh scattering cluster. As the specificity of this cluster channel relates predominantly to feature size rather than chemical composition, it is slightly less diagnostically informative than autofluorescence which reports specific biomolecules such as reduced nicotinamide adenine dinucleotide (NADH) and flavin adenine dinucleotide (FAD).[125] However, the size-dependent nature of Rayleigh scattering means this system also provides segmentation of more tissue features than would be achieved with autofluorescence imaging. In contrast to spectroscopic mapping techniques such as Raman and FTIR, while this system provides similar spatial resolution and significantly enhanced speed, its specificity is poor.[126] While this is inherent in the label-free imaging of unstained tissue sections due to their low contrast, low specificity translates clinically to a high rate of false positive diagnoses which would hinder the current diagnostic process rather than streamlining it. Therefore, as a standalone system, this hyperspectral system and segmentation software would not be adequate for assisting oral malignancy diagnoses. Optimisation of this HSI system, such as equipment substitutions to enhance scattering contributions, may yet result in diagnostically relevant segmentation. However, it would likely require combination with spectroscopic chemical mapping to provide the specificity required for a stringent diagnostic system for routine clinical use.

6 Conclusions

The preliminary investigations showed that the hyperspectral imaging system is spectrally accurate, sensitive to changes in concentration of pigmented samples and that the cluster analysis software can successfully process the large hypercubes produced and achieve meaningful segmentation from datasets with adequate spectral features. The broader wavelength source extracted more specific clustering including fibrous tissue which is well segmented, even in unstained tissue. The newly developed VB.NET software control facilitated collection of a full wavelength range hypercube, 450-750 nm with spectral step size of 1 nm, in under 30 minutes.

Two adaptations of the system were recommended as a result of these proof-of-principle investigations; to use more powerful illumination to increase SNR, and to extend the illumination wavelength range into the NIR wavelengths to exploit Rayleigh scattering for feature extraction wavelengths.

Whilst it was practical to build the system with the visible range source and gratings that were available, a great deal of the successful HSI of tissue in literature has been performed with infra-red illumination. Whilst the majority of this has been using mid infra-red (MIR) wavelengths in FTIR, near infra-red (NIR) HSI has been carried out successfully and should allow non-destructive, rapid imaging of biological tissues with a less complicated setup and greater penetration depth than MIR.[108] (FTIR was rejected as a pre-screening process due to its IR wavelength range resulting in lower diffraction-limited spatial resolution, the expense of IR detector arrays and IR-compatible optics, its long acquisition time due to scanning components, and because it requires specific sample preparation due to its sensitivity to the O-H bond in water.) The success of IR imaging of tissues is due in part to contrast provided by the enhanced optical scattering at these longer wavelengths. As these preliminary results indicated that scattering is our main contrast mechanism, in the subsequent permutation of the imaging system the balanced Halogen light source was replaced by a white light laser to extend the wavelength range into the near infrared; a highly scattering wavelength region in tissue. Additionally this powerful illumination was intended to increase the SNR to help increase the segmentation specificity.[7] Due to the slit mechanism of the monochromator, increasing spectral resolution by decreasing the slit width also results in a decrease in illumination power. To ameliorate this issue, a beam expander was installed in the light path to increase the size of the laser spot on the monochromator

[7] White light supercontinuum sources work by using nonlinear optical processes such as four wave mixing to broaden the wavelength range covered by a pump beam. They are commonly referred to as white light lasers as the wavelength components of the resulting supercontinuum have relative coherence lengths which are approximately the same as that of the incoming laser pulse and therefore the output behaves, in many ways, like a laser.

grating. This alone increased the spectral resolution meaning that the slit width was able to remain more open and allow more of the beam through, keeping illumination power high. As the beam path was being redesigned to incorporate this beam expander, reflective optics were used in place of lenses to ensure that chromatic aberrations were minimised.

To further ameliorate these chromatic aberrations we attempted to collimate the diverging light illuminating the sample by installing a fibre collimator just before the sample. To accommodate this, the liquid light guide was replaced with a single mode fibre and a ground glass diffuser was used to depolarise the light before it reached the sample. This had the additional benefit of improving the power density at the sample, increasing the SNR further.

The final modification to was to replace the sCMOS camera with an air-cooled CCD camera. The quantum efficiency of the new CCD camera in the NIR wavelength region was superior to the sCMOS, extending the useable wavelength range of the system further. An additional benefit of the new CCD camera was the substantial reduction of fixed pattern noise. Fixed pattern noise (FPN) is a constant pattern effect caused by differences in responsivity across individual detectors in an array; due to their use of sensor arrays, sCMOS cameras are prone to FPN.

This was significantly inhibiting the performance of the autofocus routine used to generate the focal plane list required for quick focus control during data collection. The grid pattern generated by the fixed pattern noise was higher in contrast than some of the sample features. Therefore, when performing auto-correlation, the software would try to detect improvement in the sharpness of this pattern with change of focal plane. Changes in focus detected by the software were often related to the fixed pattern noise rather than the z-position of the sample, leading to an extended iterative loop which would never converge on a z-stage position of optimum sample focus. CCD cameras such as the Hamamatsu ORCA do not have multiple image sensors and therefore are significantly less susceptible to fixed pattern noise. CCD camera do, however, suffer from a less problematic random "pixel FPN" due to dark current and variation across its single sensor device.

Additionally, only being able to take one field of view due to the absence of an automated x,y stage or effective FOV stitching software meant that, in many cases, the edges of the tissue would not be present in the images. Edge features are often distinctive and provide useful markers which can be used to assist in coregistration of different imaging modalities. Whilst regions of high contrast within the tissue area can vary dramatically between imaging modalities, especially between spectroscopic and optical techniques, edges usually remain clearly visible and unchanged.

The home-built cluster analysis software with overclustering was at the release stage

and we were unable to access raw output data or alter any aspect of the format in the which the results were output. To integrate these specific spectral occurence maps with Raman and widefield H&E data, new software would be required.

Part II

Improved Setup

Having demonstrated that hyperspectral images of colourless biological samples contain adequate contrast for segmentation with PCA and k-means clustering, the next unknown to explore was that of the possibility of coregistration between different imaging modalities. Comparing the HSI segmentation to the Raman segmentation will demonstrate the potential of the HSI technique as a pre-screening technique to augment the Raman diagnostic process and make it fast enough to be a practical clinical aid to diagnosis of oral cancer. It has been shown that Raman area mapping is a highly sensitive and specific oral cancer diagnostic technique. Therefore, if similar regions of interest are highlighted in the HSI imaging modality, it could be used to generate the initial segmentation of the tissue. From this map, chemical composition and health of different tissue regions could be ascertained using Raman spectroscopy. Large differences in segmentation between the HSI and the Raman screening methods would mean the techniques are too incompatible for combination and would imply that, in spite of adequate contrast, the differences in the HSI tissue data which allow for segmentation are not linked to chemical or morphological differences which are indicative of tissue health. To evaluate the possibility of multimodal comparison, a system which allows the imaging of the full tissue section, including tissue edges, was required. With no automated stage available, a software solution was employed.

Secondly, following the successful PCA and k-means segmentation of unstained tissue sections, the diagnostic relevance of the highlighted regions required verification with respect to the gold standard diagnosis by histopathological evaluation of H&E stained sections. As a new laser source was installed for this iteration of the setup which would produce data with a higher SNR and potentially affect segmentation, data for the tissue sections from Part I were re-collected and re-segmented before exploring the biological details of each region with respect to the H&E selected cancer regions. Visible tissue edges in this setup also facilitated easier multimodal comparison.

7 HSI System and Software

This setup was designed to increase SNR, and to extend the illumination wavelength range into the NIR wavelengths to exploit Rayleigh scattering for feature extraction, leading to the inclusion of a white light laser, reflective rather than lens-based optics including a reflective beam expander and a 12-bit CCD camera. Chromatic aberrations were also minimised in this setup by the switch to reflective optics, and by collimating

the previously diverging light illuminating the sample.

7.1 Setup

Specifically, this setup, shown in Figure 28, consisted of a white light laser (Fianium), a x6 reflective beam expander (Thorlabs), alignment irises and reflective focussing optics (Thorlabs), a monochromator containing a 600 l/mm, 700 nm blaze, 450-2000 nm ruled diffraction grating (Newport Cornerstone 130 1/8m), 550 µm 0.22 NA single mode fibre, 90° refective fibre collimator (Thorlabs), motorised filter flip mount (MFF101, Thorlabs) containing a 785 nm long pass filter (Semrock), x4 (0.13 NA) magnification objective (Olympus), 2" achromatic doublet lens (Thorlabs) and 16-bit air-cooled CCD camera (Hamamatsu ORCA Flash 4).

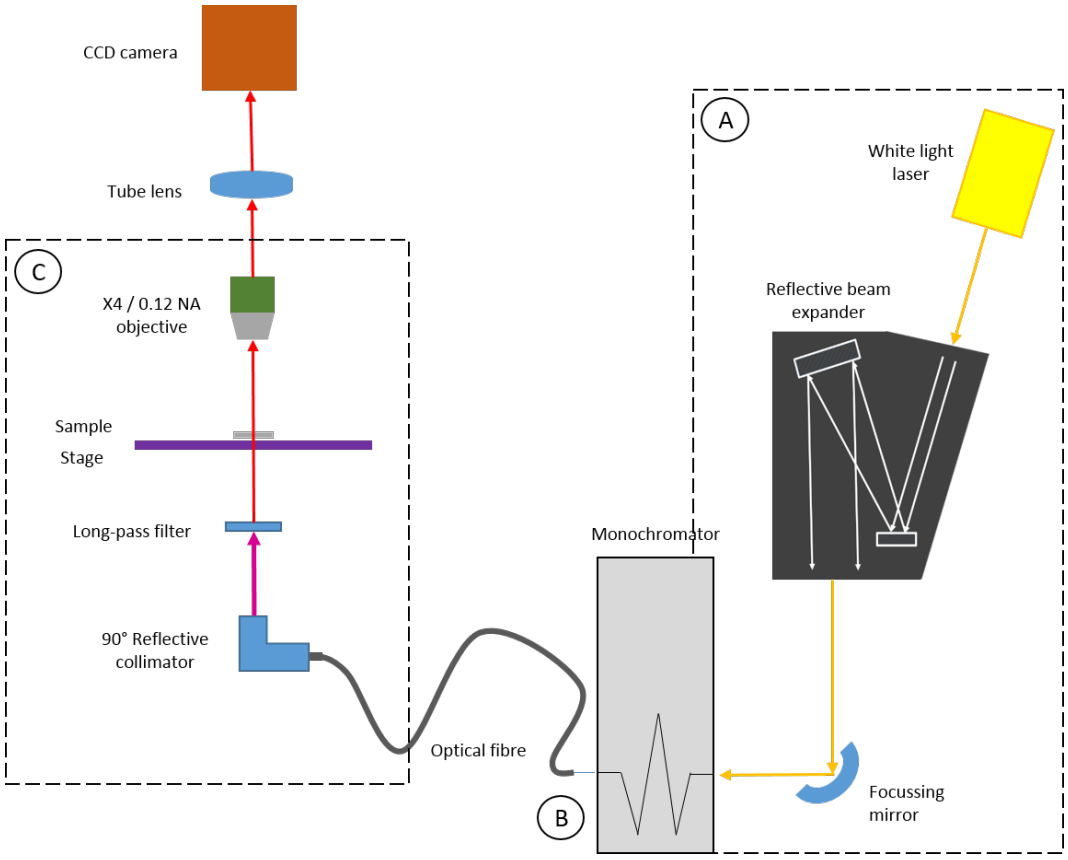

Figure 27: The light path of the improved HSI system with labelled areas A, B and C corresponding with the photographs in Figure 28.

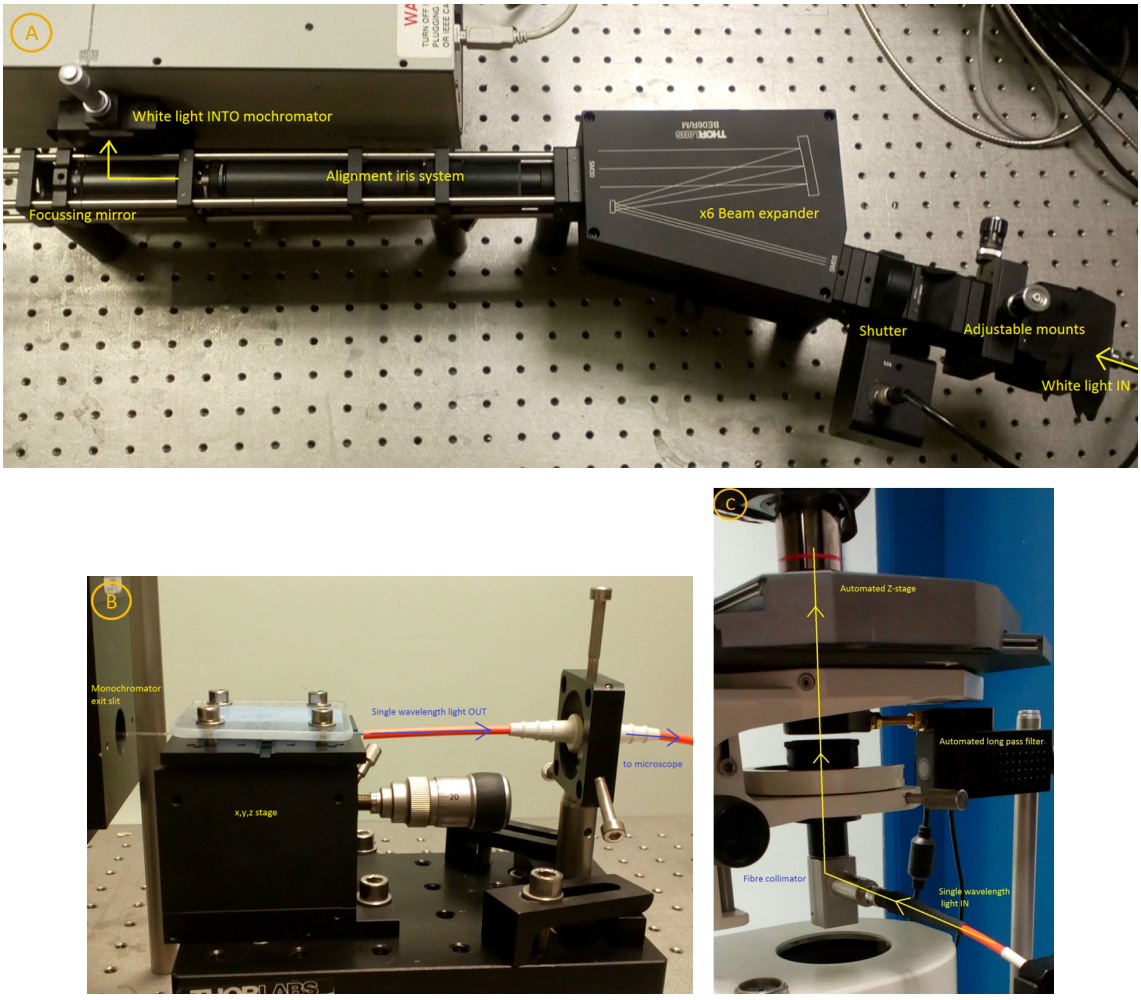

Figure 28: From top to bottom, left to right in light path order: The optics launching the white light laser into the monochromator with adjustable mounts for alignment and a reflective beam expander to improve spectral resolution. The free space fibre alignment mount for monochromatic light retrieval at the exit slit of the monochromator. The reflective fibre collimator redirecting monochromatic light upwards, through the sample and into the objective lens, via the optional long pass filter for removal of second order light. This figure corresponds with the sections of the optical path labelled A, B and C in the optical path Figure 27.

7.2 Data Collection Software

The data collection software remained unchanged from Part I and facilitated the collection of a full hypercube with user-specified wavelength inclusion, use of pre-defined exposure times at each wavelength (or autoexposure), filter flipper control for exclusion of second order light from the monochromator and concurrent dark field stack collection. Due to the newly collimated light, use of pre-defined focal positions at each

wavelength was no longer required.

For details about the background correction protocol or the removal of second order light, refer back to 3.2.

7.3 Analysis Software

7.3.1 Field of View Stitching

The hyperspectral image stacks for each tissue sample were comprised of 1 – 9 FOV which were stitched together into composite datasets immediately before analysis. The global positions of the different FOV relative to each other were approximate due to the use of a manual sample stage so the "Grid/collection stitching" ImageJ macro by Preibisch *et al* was used to calculate precise global coordinates as part of the data processing method outlined in 9.2.1.[127] "Out of memory" exceptions occurred when attempting to stitch samples comprised of more than two FOV in ImageJ so a memory-efficient stitching code was developed and added to the start of the new Matlab PCA and k-means clustering software, detailed below. This code utilised the global coordinates generated by running the aforementioned macro on just one slice of each FOV hypercube. To avoid memory issues, the mosaicking algorithm only required one hypercube to be open at a time, along with the full tissue sample hyperstack array into which the values were read, tile-by-tile.

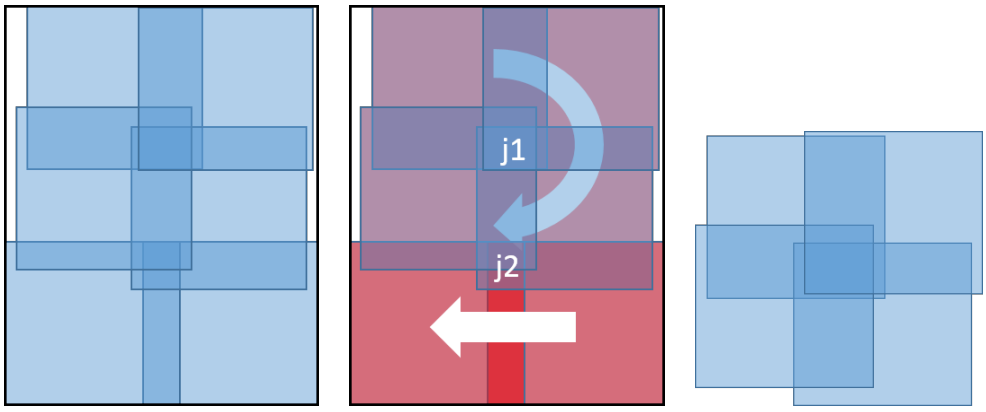

Figure 29: From left to right: An example of a 2x3 (or in Matlab, a 3x2) image mosaic with 6 FOV. Consider the two "junctions" present in this tile arrangement, j1 and j2. For j1, from top left tile, step clockwise around the junction opening the image stack for that tile and inputting their grayscale values. Repeat for j2, but ignoring the first two tiles as they are members of j1. In areas of tile overlap, the maximum value is selected from contributing tiles.

This was achieved by filling the values from each tile, moving in a clockwise fashion around each "junction" (Figure 29), closing individual tile stacks when their values were

fully transferred to the array which described the complete sample. Testing of different parameters in the Preibisch *et al* mosaicking macro showed that blending overlapping tiles by taking the maximum available pixel value tended to produce sharper features than linear blending of tiles, as demonstrated in Figure 30.

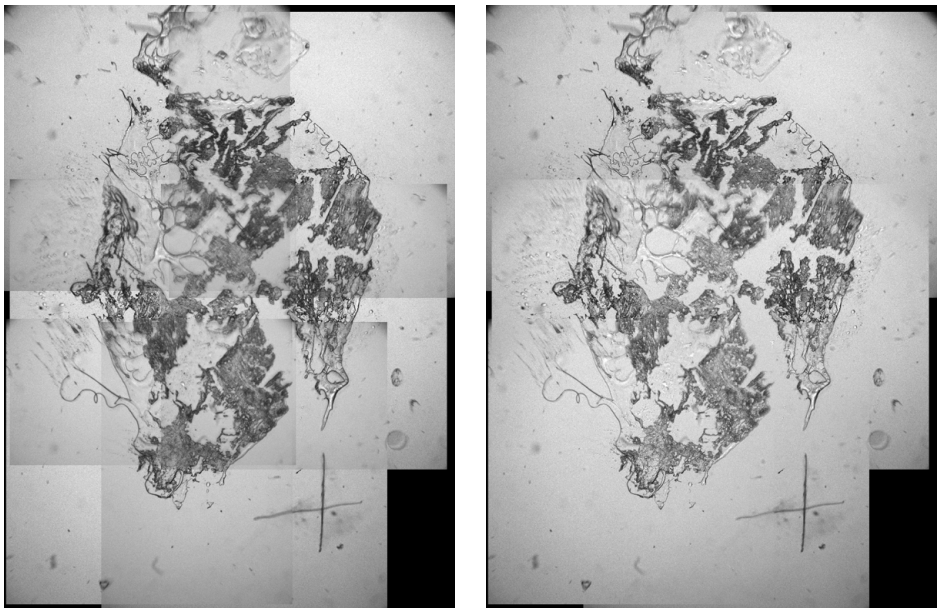

Figure 30: An example of mosaicking performed utilising the linear blending method (left) and the maximum method (right). The former takes the mean of the values which contribute to areas of tile overlap, and the latter picks the maximum value from contributing tiles.

Therefore, in the Matlab tiling code, pixel values in areas where tiles overlapped were filled in and updated iteratively. The maximum available pixel value was evaluated, and retained, each time another "option" was provided by a new hyperstack tile being opened.

The home-built cluster analysis software with overclustering used in Part I was at the release stage. The output graphic from the software affected the aspect ratio of the image which, in combination with the low spatial resolution of the Raman with the spatial binning required for practical speed, led to distortion of highlighted features which made them extremely difficult to match with features in other imaging modalites. Being unable to access and manipulate the raw output data meant that coregistering the Raman k-means cluster maps with other imaging modalities for comparison was extremely difficult. Therefore a very similar PCA dimensionality reduction and k-means clustering programme was developed in Matlab. Two main factors influenced the selection of these analysis techniques: the main contrast mechanism being scattering and the practicality of not requiring prior information about the sample (clinical input)

or the pathology of healthy and malignant oral tissue (training data). Re-writing this programme allowed this output to be tailored for comparison with other modalities: saved as tiff images and matched in pixel dimension to the input data.

7.3.2 PCA Noise Filtering

Principle component analysis (PCA) was performed on the complete mosaicked datasets, according to the method outlined in step 1 of 3.3.1, before clustering as it improves the performance of the clustering by eliminating some of the noise in the data. It does this by reducing the data to its fewest possible dimensions by translating the data into a new feature space which fully describes the data with linear combinations of the pre-existing features (principal components). The number of principal components required to fully describe the data can be identified by plotting the variance of the data explained by each component; the curve will flatten when the addition of more principal components adds little value (explains little additional variance of the data). Usually these "additional" principal components are noise and reconstruction of the data without them increases its SNR, though this process does risk accidental exclusion of important information.

7.3.3 K-Means Segmentation

Overclustering (step 2), and associated optimal cluster number selection (step 3) of the analysis software outlined in 3.3.1, were replaced by the following (standard) k-means clustering protocol which converges significantly more quickly.

1. **K-means Clustering**

   (a) The k cluster centroids (mean spectrum of each cluster) were initialised using Monte Carlo. How similar each spectrum was to this centroid was then calculated in terms of the Pearson Correlation Coefficient starting at k = 2.

   (b) Each spectrum was then assigned to the cluster with the most similar centroid, shown by the largest PCC value (r_{max}).

   (c) All of the spectra assigned to each cluster were then averaged to produce new centroid values for each cluster.

   (d) This process was repeated until the spectra assigned to each cluster remained constant.

The cluster assignation (step 4) outlined in 3.3.1 was then performed. This was the approach used in Parts II and III.

K-means clustering has been used to good effect in similar investigations and produces segmentation with maximum cluster separation without prior knowledge of the

separable characteristics of the data, though the initial randomised seeding of these centroids does decrease repeatability.[89] A fixed cluster number was therefore selected (discussed in 9.3.2) and these groups within the data were assumed to correspond to different components within the tissue sample, for example healthy and malignant epithelial tissue. A randomised colour map of where each of these clusters occurs in the input image was produced as the output.

8 Raman System and Software

8.1 Raman Setup

The Renishaw InVia Raman spectrometer (Renishaw, Gloucestershire, UK) was used to collect Raman area maps of the tissue with Renishaw's WiRE 3.4 software. The spectrometer emits a laser line which is focussed by reflective optics onto the sample via a microscope. Both Rayleigh and Raman scattered light is emitted from the sample, with the former being filtered out and discarded and the latter being directed onto an array of finely spaced lines on a reflective surface, a grating, which disperses the light into its constituent wavelengths onto a CCD for recording. The range and precision of this spectrum depends on the lines/mm of the grating and is chosen based on the wavenumber of important spectral features, which relate to the chemical bonds present in the sample. Additionally, the wavenumber range will affect the signal-to-noise ratio (SNR) of the resulting Raman spectra so the intensity of useful spectral features must also be considered when making this decision. The probability of Raman scattering is significantly lower than Rayleigh scattering and therefore to produce spectra with an adequate SNR requires pointwise scanning of the sample which is extremely time consuming. Renishaw have increased the speed of high SNR area scan acquisitions by utilising a cylindrical lens to distribute a high power laser line across the sample surface. This technique and software accommodations, known together as Streamline™, allows several point spectra along the laser line to be collected simultaneously without depreciating spatial resolution in x or y and decreases the total time required to scan the laser across the full surface of the sample.[128]

8.2 Analysis Software

The same home-built PCA and k-means clustering programme used to analyse the HSI data was used to analyse the Raman spectral maps. For details, see 12.3.1. Tiling was not required for the Raman dataset as the full tissue area is covered by point-scanning of the user-defined area; usually 0.5-1.5 cm^2.

9 Materials and Methods

The materials used in in Part II were the same as those used in the unstained tissue investigation in Part I with the hope that improvements between the two systems and analysis methods would be clearly demonstrable. The methods in Part II are also similar to the unstained tissue investigation in Part I, but with multiple fields of view and an attempt at coregistration and comparison between different imaging modalities.

9.1 Mixed Healthy and Malignant Oral Tissue Sections

9.1.1 Sourcing

Patient tissue samples and data were provided by Guy's & St Thomas' Head & Neck Biobank – part of the KHP Cancer Biobank, which is supported by the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award and Guy's & St Thomas' NHS Foundation Trust. The biobank has ethical approval from the East of England - Cambridge East Research Ethics Committee (18/EE/0025).

This study utilised three healthy and three malignant oral tissue samples which were randomly selected from the sample set of Dr. Manickavasagam to ensure a range of anatomical sites and malignancies were represented. Imaging was undertaken "semi-blind" in that I was aware of the diagnosis for each tissue section (in order to select appropriate samples) but not what features on the slide led to each diagnosis.

9.1.2 Preparation

Two subsequent (or as close as possible) tissue sections were cut from each of the listed tissue samples. One 4 µm section in frozen section fixative was stained with H&E on a glass microscope slide and topped with a glass coverslip, for histopathological review, and one 10 µm frozen section was sandwiched between two round Raman-grade calcium fluoride slides for hyperspectral imaging and Raman spectroscopy. For full protocols, please refer back to 4.2.2.

9.2 Data Collection

9.2.1 Hyperspectral Imaging

Hyperspectral image stacks were collected using the software detailed in 3.2 with a x4, 0.13 NA objective lens (Olympus) across a wavelength range of 450-950 nm in increments of 5 nm producing a 97 slice hyperstack for each field of view (FOV). Optimal camera exposure times and stage focal positions for each wavelength were read in from

two pre-collected text files throughout the hyperstack collection process. At 785 nm, the long pass filter was inserted into the optical path to filter out second order light. Collection of corresponding dark field stacks was interleaved with raw image collection by triggering the automated shutter of the monochromator and taking an unilluminated, but otherwise identical, image at each wavelength to capture and correct for background light, dead camera pixels and other dark-visible imaging artefacts. This resulted in a total, unoptimised run time of approximately 17 minutes.

For full data collection instructions and reference to the software GUI, please refer back to 4.3.1.

Hyperspectral Data Processing

Shift Correction

The microscope stage was prone to shifting slightly, and inconsistently, in x and y over time which meant that alignment errors were present in several hyperstacks. As these position shifts did not happen in a fixed direction or at a constant rate, standard ImageJ drift correction plugins performed poorly in correcting this. A shift correction macro was developed which applied a user-specified translation to a subset of stack slices. The required translation was identified by marking one linear image feature, both in a wavelength slice before the shift, and in the first shifted wavelength slice, with the line drawing tool in ImageJ. These two selections were then added to the ROI manager and their relative positions retrieved from their automatically generated names, which take the format "z position-y position-x position". Subtraction of one from the other produces the translation required to correct the shift using the aforementioned macro.

Brightness Fine Correction

A simple brightness correction macro was also developed in Image J which utilised a user-selected reference brightness area, selected by a square region of interest (ROI), to correct the relative brightness of each wavelength slice of the hyperstack. This was performed on each stack after shift-correction, to ensure intensity measurements were being compared across the same physical area in each slice.

Flat-Field Correction

Originally, a flat field stack was collected once per objective, by taking an exposure and focus-corrected hyperstack across the full wavelength range with the shutter open but with no sample on the stage, with the intention of performing flat-field correction

after the data was collected. This allows the capture and correction of inhomogeneous illumination, illumination power variation across the wavelength range, specks of dust trapped in the optical pathway and other light-visible imaging artefacts. In practice, however, this approach was hindered by interference effects which presented as concentric "squircles".[8] The distance between the rings related to the illumination wavelength but unfortunately, while the separation of these concentric rings were fixed, their positions were not. They streamed outwards from the centrepoint of the squircle over time. Broadly, "fringe shifting" occurs when the relative phase of component sources changes. White light lasers are usually considered to be semi-coherent because, whilst they produce spatially coherent light, it cannot be spectrally coherent as it is not monochromatic. Additionally, the measured spectral resolution limit of this laser used in conjunction with the Cornerstone monochromator is approximately 2 nm. This not-quite-monochromatic or coherent light provides some of the multiple potential mechanisms by which this fringe shifting effect could occur. The most likely explanations are beating between two modes of the source, a reflection in the cavity, interference between different modes of the fibre or interference between dominant wavelengths of light which are constantly changing due to instability in the laser. This interference, combined with the slight flicker of the white light laser meant that when dividing the dark field-corrected stack through by the flat-field stack, unwanted variations in brightness were added rather than removed. This approach was therefore abandoned in favour of the simpler ImageJ brightness correction macro detailed above, which was initially created for fine correction of minor random source brightness fluctuations.

Field of View Stitching

Multiple fields of view were collected from the top left, snaking right and down. Each field of view was flat-field, shift and brightness-corrected as above, and they were stitched together using the "Grid/collection stitching" ImageJ macro by Preibisch *et al.*[127] Diverging light passing through the sample led to depreciating focus towards the outer edges of the FOV. As a result, a minimum of a 40% area overlap between neighbouring fields of view was required for accurate stitching. If a pre-screening imaging techique which relied on fluorescence had been selected for this application, this generous overlap area would have risked photobleaching. The parameters that were most commonly used for the stitching process are shown in Figure 31, with small variations in percentage overlap dictated by trial and error.

[8] A squircle is an intermediate shape beween a square and a circle.

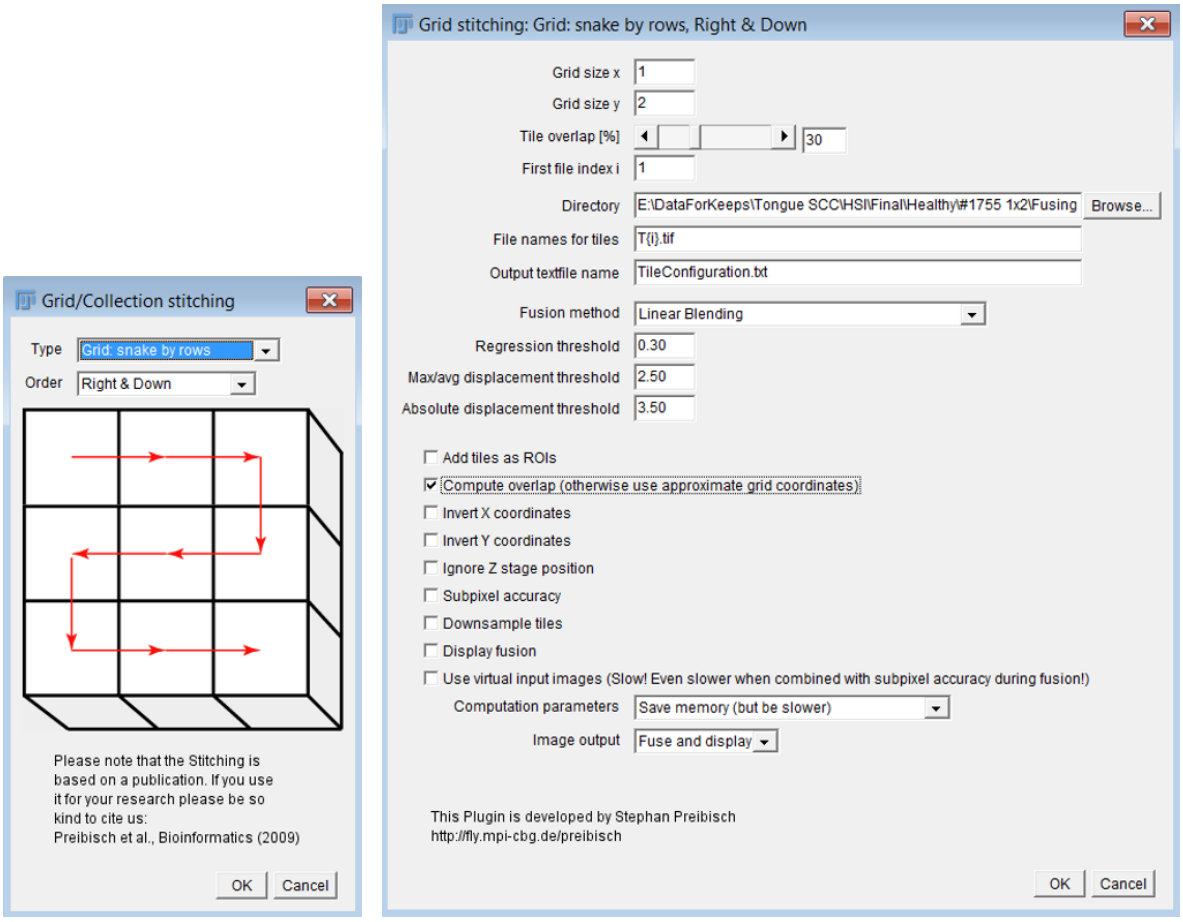

Figure 31: The two GUIs from the Preibisch *et al* grid/collection stitching macro, showing adjustable parameters and the values or options most commonly used in this study.

Initially the intention was to stitch the hyperstacks for each tissue sample together, in full, to produce one very large 3D dataset for simple display and analysis of each sample dataset. Unfortunately, the large percentage overlap required due to the unflat field rendered a large proportion of each FOV unusable. This meant that up to 9 FOV were needed for each sample, each with 97 slices, producing memory errors in ImageJ when full stitching was attempted. Instead, a more economical stitching method was developed and built into the analysis software (Figure 29).

9.2.2 Raman Spectroscopy

An RGB image montage of the tissue and surrounding area was taken by the InVia Raman spectrometer with a x20 (0.45 NA) objective, allowing manual selection of the desired area for Raman scanning. After which, the Raman spectrometer was run in Streamline™mode to collect an area map of Raman spectra. A 785 nm diode laser was used in conjunction with a 600 l/mm grating. This excitation wavelength is ideal

for tissue as it is non-destructive, results in less tissue autofluorescence than lower wavelength sources and produces data with a higher SNR than higher wavelength IR sources. The low line density grating means that spectra are collected over a broader wavenumber range (~ 981 – 1719 /cm) with a lower spectral resolution (1.9 /cm per pixel). This is well suited to the broad fingerprint regions exhibited by many tissue components such as collagen.[129] These peaks were also used to inform the central wavelength of the spectral range, which was chosen to be 1355 /cm. 100% laser power with a 2 s exposure time was found to be an acceptable balance between SNR and run time, having set an upper limit of a 16 hour full tissue section scan time. The spatial resolution (step size) of the area scan was 50 µm in the x-direction and 13.7 µm in the y-direction, with the WiRE 3.4 software Ybin parameter for Streamline area scanning set to 5 based on optimisation work performed by Manickavasagam *et al.*[120]

9.2.3 Histopathology

RGB images of the H&E stained tissue sections were collected with a Olympus BX60 Biological light microscope x10, 0.25 NA objective (Olympus), illuminated with a tungsten halogen source and captured with a Basler Pilot piA1600-35gc area scan camera.

Histopathology Data Processing

Field of View Stitching

The fields of view were mosaicked using the Preibisch *et al* stitching macro detailed in 9.2.1, using the "Maximum" method and a 20-30% percentage overlap determined by trial and error.[127]

9.3 Assessment of Tissue

After collection and processing, the gold standard reference data for each tissue section was one RGB image, the Raman data was a .wdf file which stores the Raman spectra of the area scan as a list of intensity values and the hyperspectral data was a stack of 97 grayscale images. To compare these very different modalities directly, and to assess their relative potential for identifying cancerous tissue regions, they first needed to be transformed into similar formats and then segmented. For simple and easy-to-interpret comparison, the desired format for all three modalities was one image with cancerous regions highlighted by an area outline or fill.

9.3.1 Histopathologist Segmentation of H&E Data

The RGB images of the H&E stained tissue sections were evaluated by two expert histopathologists. In ImageJ, regions of interest were manually selected in the images of the three tissue samples selected from the KCL biobank library of cancerous oral tissue (#8460, #8883, #9120). In the case of the three healthy tissue sections (#232, #6111, #4984), no regions of interest were selected as samples in which precancerous changes or cancer cells were identified were excluded from this group. Selections were made using the following protocol.

1. Open .tif RGB image of H&E stained tissue.
2. Select "freehand selection" tool from ImageJ toolbar.
3. Draw around cancerous region with precision on the order of a few cells.
   (a) To zoom, hold cursor over desired area and press "Ctrl" + "shift" + "+".
   (b) To view all ROIs simultaneously select the "Show all" checkbox in the ROI manager.
4. To add first selection to ROI manager by clicking "Edit > Selection > Add to manager" in ImageJ toolbar.
5. To add subsequent selections to ROI manager, click "Add" at the bottom of the ROI manager or press the "t" key.
6. To save ROIs, select all ROIs in ROI manager and click "More > Save" in ROI manager.

9.3.2 Segmentation of HSI Data

The segmentation of the HSI data was undertaken using the home built Matlab PCA and k-means clustering software described in 7.3. The two variables which can be altered by the user in this software are component number and cluster number; the first of which related to the PCA and the second to the k-means clustering. The component number is the number of dimensions which are selected to express the image information. Cluster number is the number of centroids used to seed groups of the data, classified by certain similarities.

Ordinarily, when investigating known samples with a known technique, knowledge of the expected SNR can inform the choice of fixed principal component number. Utilising all of the generated components, for example, would include all of the signal and noise in a dataset, which could obscure useful segmentation. On the other hand, using too

few of the components would lead to data with selections which are less specific and which may not contain more subtle separations. Sections of tissue containing absorbant compounds such as red blood cells may be separated out from the rest due to their strong contrast but hypothetical areas of slightly increased scattering from increased nuclear-to-cytoplasmic ratios due to cancer, for example, may not be separated. In this exploratory work, more was known about the desired tolerance of the outcome than the quality of the data being used to generate the segmentation. Therefore the component number selected was based on the turning point of the cumulative plot of percentage data variance described by each component, shown in Figure 32; to accommodate differences between samples, this was set at 7.

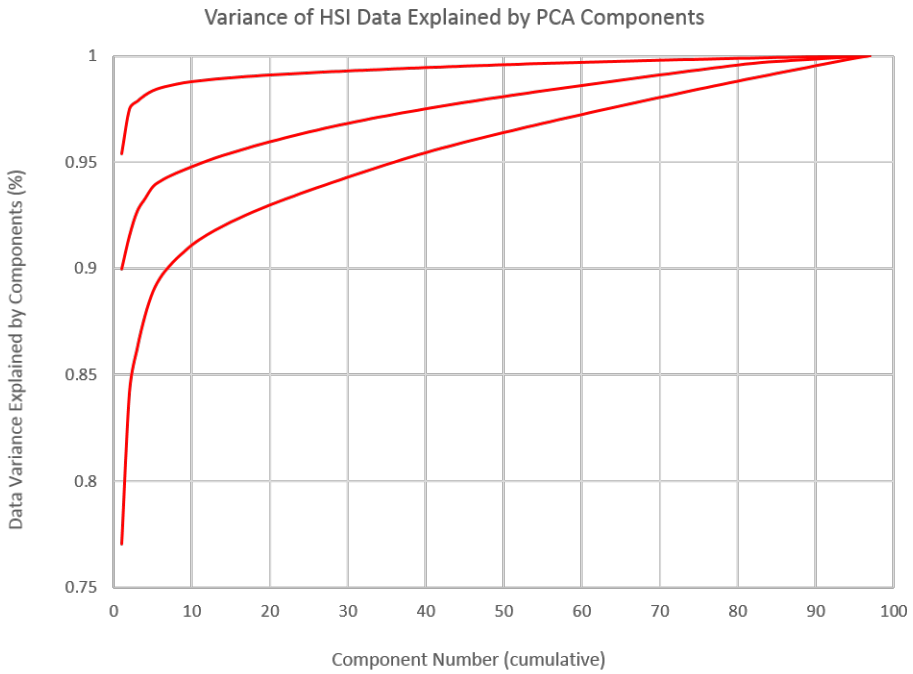

Figure 32: A graph of the variance proportion of the HSI data of the three malignant oral tissue samples explained by each component as determined by the Matlab software detailed in 7.3.2.

The selection of a user-defined cluster number for the k-means clustering requires a little more prior knowledge of the sample. Often cluster number is manually selected on a case-by-case basis by a pathologist who identifies the number of different tissue components in the sample. Alternatively multiple segmentation maps are produced, each with a different cluster number, and the pathologist will identify which map and associated cluster number most accurately describes the tissue. The k-means overclustering software outlined in 3.3 worked well with the single field of view datasets collected in Part I, however, most datasets collected in Part II were several times larger. The computing time required to run PCA using this overclustering technique for cluster

number selection on these significantly larger datasets would be impractical in the clinical setting for which this technique is intended. Instead, to facilitate the use of regular k-means clustering, a cluster number was chosen for each sample. This was based on the assumption of one background cluster per field of view tile and six main tissue components found in cancerous oral tissue sections:

1. Healthy epithelium
2. Unhealthy epithelium
3. Muscle
4. Inflammatory cells
5. Connective type tissue (mostly collagen, keratin, walls of ducts etc.)
6. Adipose

9.3.3 Segmentation of Raman Data

The Raman streamline map data, being a .wdf file type, was rearranged into the standard x,y,z(λ) format by an open source Matlab programme released by Renishaw called renishaw_Wire.m. The home-built PCA and k-means clustering software from 7.3 was used to denoise, and then segment, the Raman data. The component number generation was automated so that 95% of the data variance was explained and the cluster number was set to 7, according to the rule outlined above. Raman has been shown to detect subtle differences in tissues of different types and health statuses, such as cancer and inflammation, and is therefore sensitive enough to successfully segment the same tissue with a higher cluster number than is possible with the HSI data. However, reducing the cluster number should reduce the separation to the 7 most clearly delineated tissue types or features, with the hope being that these are the same 7 clusters which can be identified in the HSI data.

9.4 Cross-Comparison

Using the open source image processing software GNU Image Manipulation Program (GIMP), the coloured maps of the three imaging modalities were overlaid and transformed to match each other to allow for clear comparison of the regions of interest selected by each technique, as shown in Figure 33.

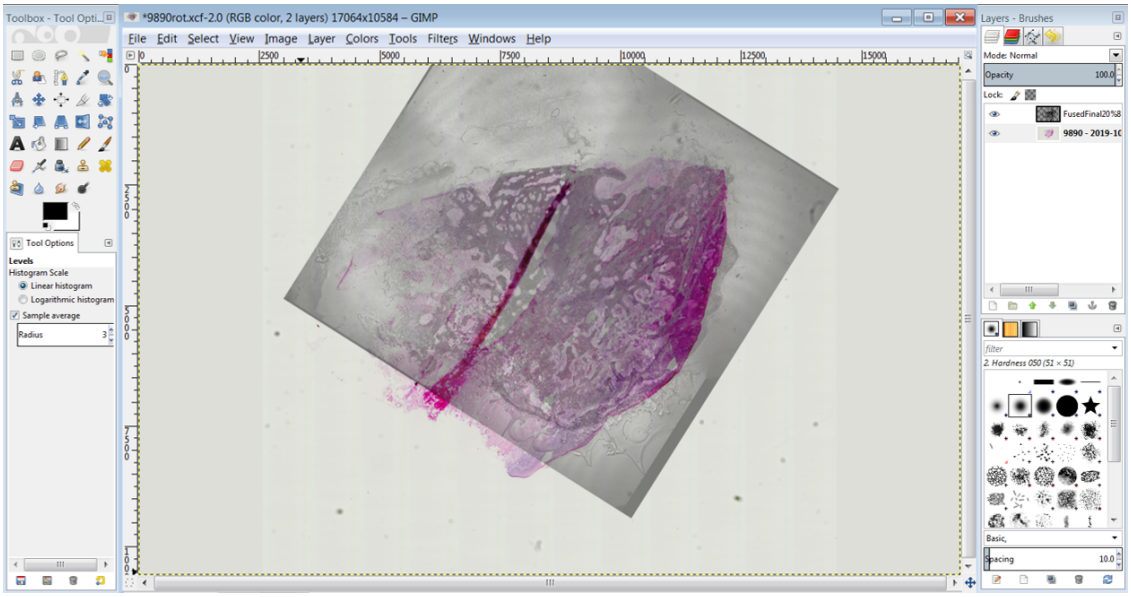

Figure 33: Printscreen of GIMP overlaying process.

This process followed the method outlined below:

1. Open the RGB tiff image of the H&E stained tissue section.
2. Identify its most solid edge, which is often the epithelial edge. This is less likely to have lost fragmented pieces of tissue, changing its shape.
3. Open the PCA map of the HSI or Raman data stack as a layer by clicking "File > Open as layer".
4. Use the sliding opacity scale in the "Layers" sidebar, decrease the opacity of the PCA layer so that the H&E stained tissue is visible through it
   (a) a. If necessary, click " > Levels > Color Balance" and alter the sliding scales to increase the contrast and brightness of the H&E image.
5. Identify the matching solid tissue edge in the PCA image and align it with that of the H&E stained tissue in the RGB reference image.
6. Using the transform tools available on right-clicking the PCA layer in the layer sidebar, alter the image orientation and scale iteratively until the two tissue images match.
7. Save the layer separately to the H&E reference image.

This process allowed spatially matched images from each modality to be produced. Similarities between regions segmented from each technique were then easily identified by eye.

10 Results and Discussion

This study was performed on three healthy tissue samples, two malignant samples carrying diagnoses of adenoid cystic carcinoma and squamous cell carcinoma and one benign tumour diagnosed as pleomorphic adenoma. Their various oral cavity sites can be found in Table 2.

10.1 Healthy Samples

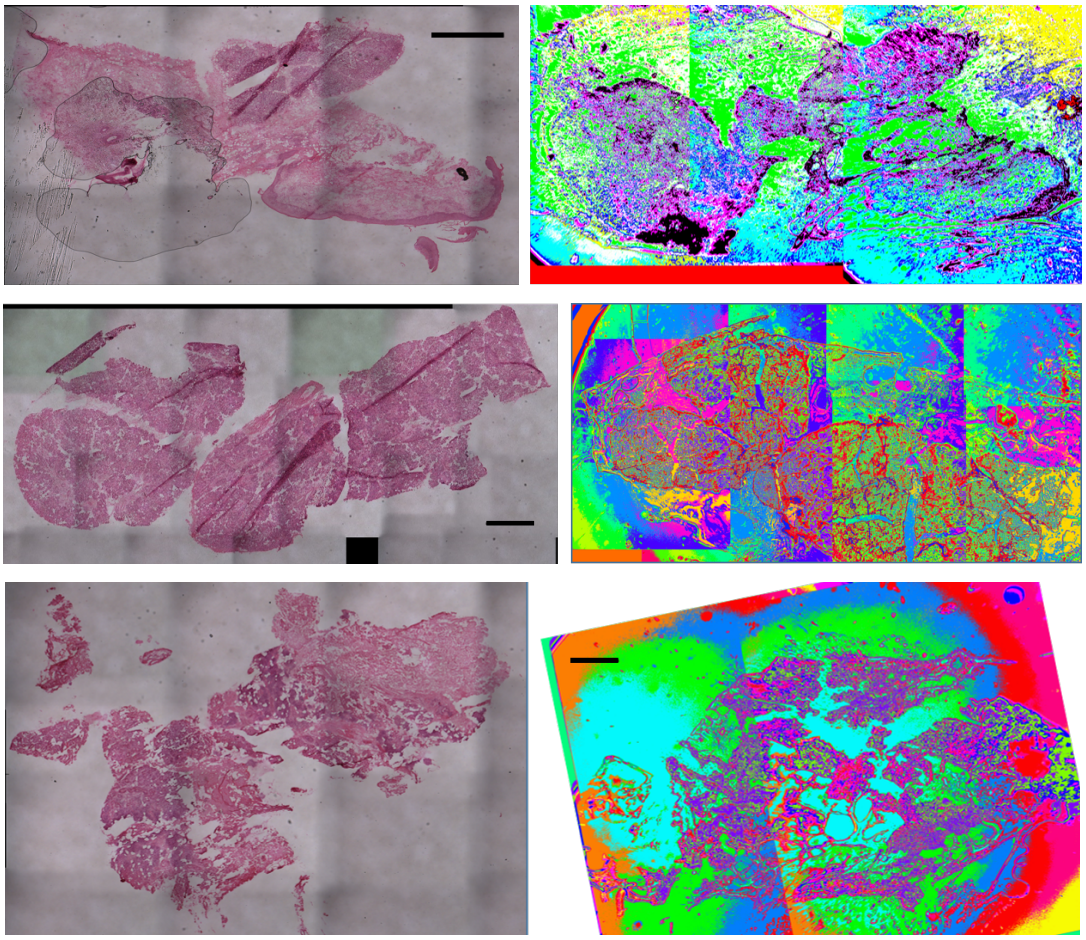

Figure 34: Healthy tissue samples #6111, #232 and #4984 (top to bottom) in pairs (left to right) of H&E stained section RGB image and coregistered, PCA dimensionality-reduced, k-means clustered HSI of subsequent tissue section. Black scale bars represent 1000 µm.

Both sections of sample #4984 were dissimilarly fragmented rendering accurate coregistration impossible (though a best attempt is included in Figure 34). To a lesser extent, sectioning artefacts have also impacted sections #6111 and #232, with coregistration being successful only on the right hand side of the former, and the left hand

side of the latter, where matching tissue regions are self-evident. This unfortunately means that the data on the left of the image of #6111 and the right of the image of #232 is redundant, having lost the option of comparison with the H&E images of the associated stained sections. In all sections, a tiling artefact has substantially impacted the segmentation. This is seen as a vertical or horizontal (before translation for coregistration purposes) line where the colour of a segment changes abruptly from one side of it to the other (Figure 35).

This occurs because the background of the image is a large, relatively uniform area which is therefore easy to classify into a single component or small number of components. However, slightly different intensity levels between the different images being tiled mean that background originating from different images is assigned to different components. For clinical and pre-clinical use, the analysis could be augmented by using areas of overlap to identify the spectrum of the background cluster across multiple fields of view, and then use that to assign a single background segment.

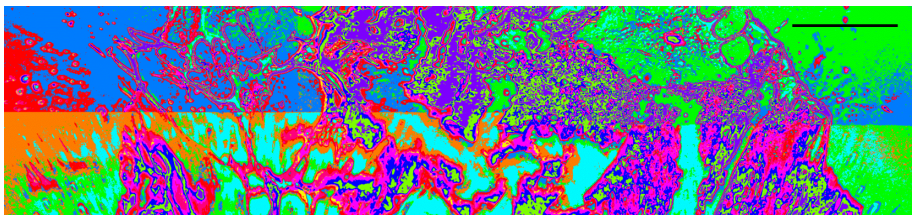

Figure 35: A magnified view of an area exhibiting a tiling artefact. Black scale bar represents 500 µm.

While the tiling artefacts and large and highly varying, squircular background signal (detailed in 9.2.1) obscure most internal tissue feature segmentation, the tissue is clearly delineated from the background and the OCT in the clustered HSI. (The distinct black segment present in sample #6111, predominantly found at the bottom left of the sample, is suspected to be residual ink from markings made during the tissue biopsy process.)

10.2 Malignant Samples

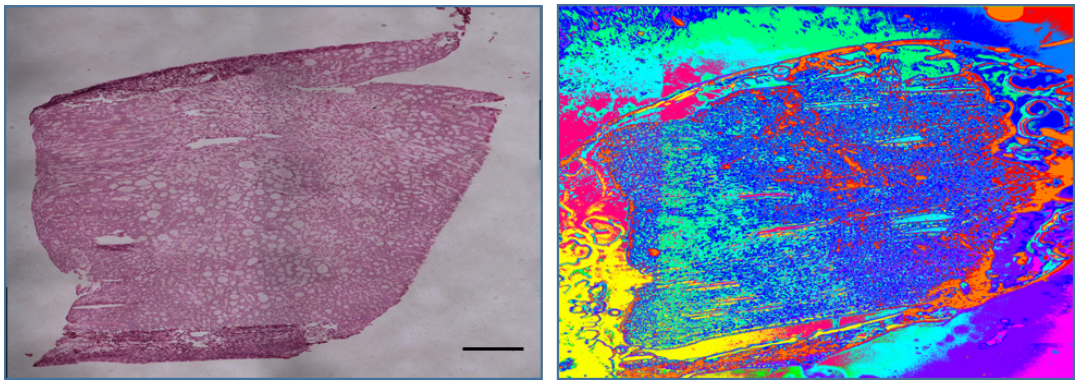

Figure 36: Parotid gland adenoid cystic carcinoma #9120 H&E stained section RGB image and coregistered, PCA dimensionality-reduced, k-means clustered HSI of subsequent tissue section (left to right). Black scale bar represents 1000 µm.

Adenoid cystic carcinoma is a malignant secretory epithelial tumour with a deceptively benign histologic appearance which begins in glandular tissue; in this case, the parotid salivary gland.[130] This sample shows the most common tumour growth pattern, cribriform, which is a sieve-like collection of small islands or sheets of cancer cells with cyst-like spaces which can be seen in Figure 36. In H&E stained sections, these cyst-like acellular cavities containing some secretions which stain light pink with haematoxylin, and the cancer cells themselves, comprise the rest of the sample and are characterised by enlarged purple stained nuclei which fill a large area of the cell. As there is no solid tumour body, and the cancer is instead interwoven with the cyst-like spaces throughout the section, it was not possible for histopathologists to select regions of malignancy.

Evaluating the performance of the segmentation is therefore more challenging. One histopathologist commented that the centre of the sample is more cystic and the density of the tumour appears to increase towards the section edges, especially the left, of this section. Errors introduced by tiling multiple fields of view are evident in this sample (see Figure 36), with the pale green segment selected with increasing density to the left of the tissue area ending abruptly with a vertical tile edge. This segment is not present along the right hand edge of the section as the histopathologist indicated and this pale green area correlates with with an suggested area of increased cancer density if we take the tiling artefact into account. Alternatively, the absence of the pale reen segment along the right hand edge may be due to a real difference between the H&E stained tissue section and its paired, but not identical, unstained section for HSI.

However, it could be artefactual and caused by blurring which is present in areas of tile overlap in this sample. The correlation function which the ImageJ mosaicking macro relies on performs less reliably in samples with multiple feature edges spread homogeneously across the samples, such as the cyst-like spaces. At the top of the tissue

it's possible to identify the edges of the central field of view tile, of which there are 6 (3 columns and 2 rows). A substantial increase in the area covered by the pale green segment occurs from the right to the left of this tile alone, suggesting that this is linked to a real change in tissue characteristics rather than a tiling artefact.

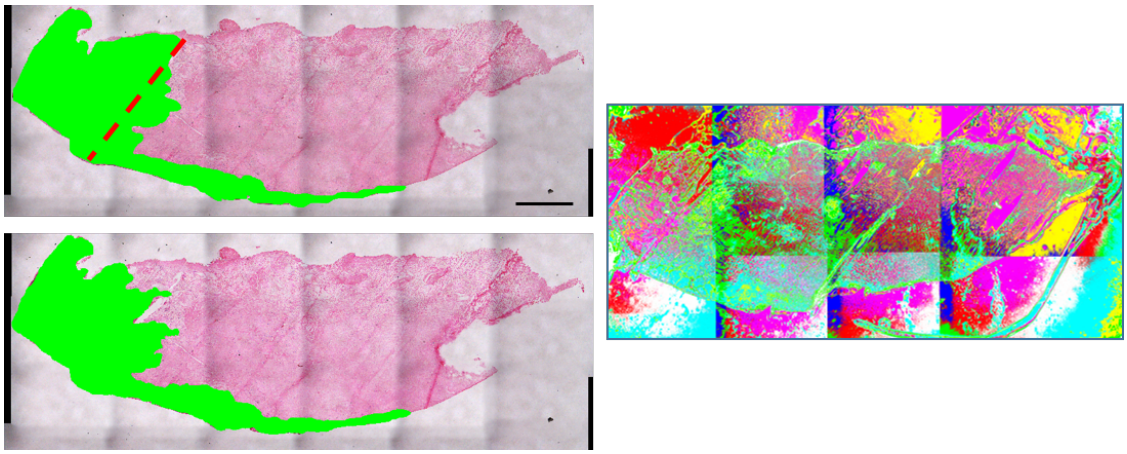

Figure 37: Parotid gland pleomorphic adenoma #8883 H&E stained section RGB image (left). The green area is an area of cellular pleomorphic adenoma independently demarcated on separate but identical images by two histopathologists, and the dotted red line indicates the position of a series of holes. The coregistered, PCA dimensionality-reduced, k-means clustered HSI of subsequent tissue section (right) which appears to have severed, and then fragmented, along the line of holes marked with the red dotted line in the H&E section. Black scale bar represents 1000 µm.

Unfortunately this sample was mis-filed and is, in fact, a benign tumour, though it does have high potential for malignant transformation to carcinoma ex-pleomorphic adenoma which occur in approximately 8.5% of cases, though estimates of this progression vary and increase dramatically over time.[131] The cellular region appears to have been segmented (green), though other components have been grouped with it including several small areas of keratin, most of the outer tissue edges and areas where the tissue section has folded (see Figure 37). All of these tissue features, except the tissue edges, would be expected to have increased scattering compared to the surrounding areas which are comprised of tumour fibrous capsule which accounts for their separation. This fibrous capsule typically envelops these tumours *in vivo* and is a natural plane of cleavage, reflecting the risk of resection at the capsule margin with any tear leading to tumour spillage and field adulteration. Unfortunately the extremely dominant varying background signal claims most of the fixed number of clusters in the k-means clustering software. It may be the case that, once this background contribution is minimised or flattened, or with a higher cluster number input for the segmentation, these tissue features could be identified as different and assigned to their own clusters. Though inclusion of a benign sample was not intentional, it has provided additional insight

into the capabilities of the HSI system's diagnostic capacity. While the cellular region, keratin and tissue folds have increased scattering compared to the surrounding tissue, this scattering difference is very likely to be less marked than between healthy tissue and cancer. The technique therefore holds promise for successful segmentation of oral cancer.

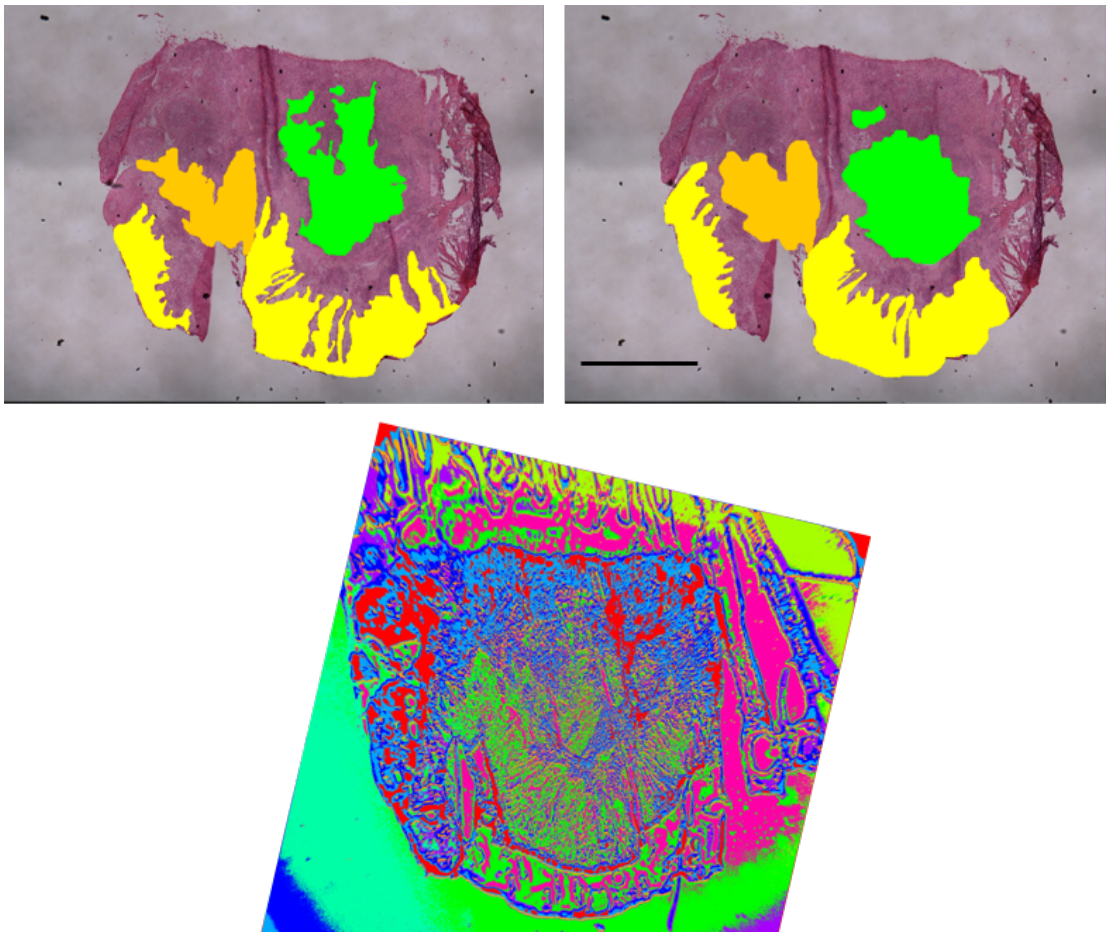

Figure 38: Oral squamous cell carcinoma sample #8460 H&E stained section RGB image with cancer (green), dysplasia (orange) and hyperplasia (yellow) independently demarcated on separate but identical images by two histopathologists (top) and coregistered, PCA dimensionality-reduced, k-means clustered HSI of subsequent tissue section (bottom). Black scale bar represents 1000 µm.

Sample number #8460 was substantially smaller than the other malignant samples in this study and it was therefore possible to capture the HSI data for the full tissue slice in one field of view. Due to the absence of FOV tiling artefacts, this single field of view dataset reinforces the assertion that the internal tissue area segmentation may be related to biological features.

Hyperplasia (demarcated in yellow on the RGB images of the H&E stained tissue section in Figure 38), results from upregulated cell proliferation and causes increased

cell numbers, dysplasia (orange) is when cell proliferation begins to be unregulated and they appear abnormal but do not yet meet the criteria for malignant neoplasia (cancer). Both of these changes predispose the cells to becoming cancerous, but do not guarantee it, and cells need not progress through these two stages to become cancerous.

The areas of cancer, dysplasia and hyperplasia in the segmented HSI image all appear to have been selected in green. This implies that, to some extent, unhealthy oral tissue undergoes changes which are reflected similarly in their respective elastic scattering spectra. The common pathological features of hyperplasia, dysplasia and cancer are increasing cell density as well as increasing size and density of subcellular features including nuclei and mitochondria. Elastic scattering can be sensitive to all of these spatial alterations, though it depends on the angle of observation. Wavelength dependent scattering alterations, at a fixed angle, are predominantly sensitive to the average size of scatterers.[132]

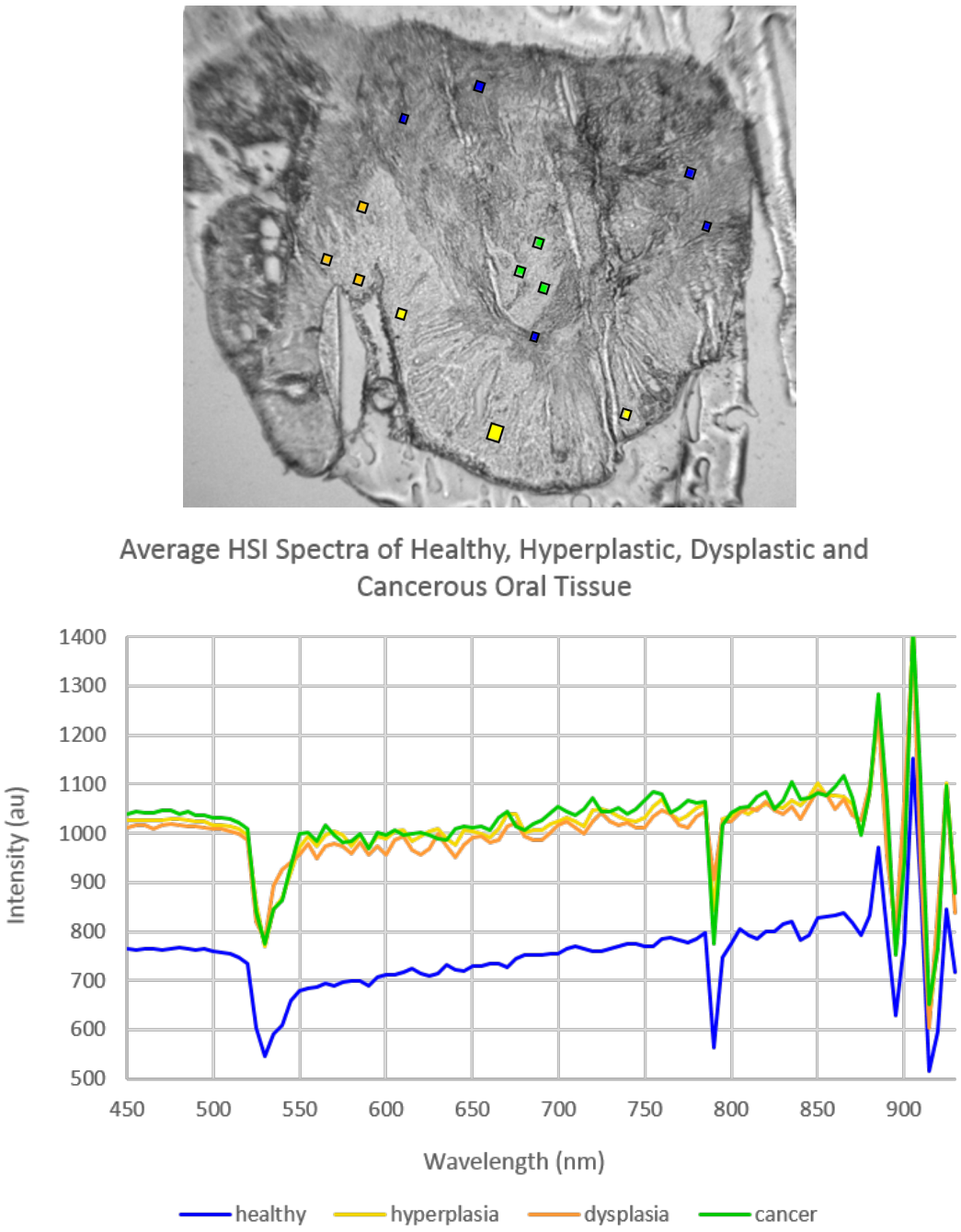

Figure 39: A single slice (550 nm illumination) of the HSI stack of oral squamous cell carcinoma sample #8460 with sample regions of healthy, hyperplastic, dysplastic and cancerous tissue highlighted in blue, yellow, orange and green (respectively) and graph with averaged spectra from each area.

Regions of interest were manually selected within the pathologist-selected areas of hyperplasia, dysplasia, cancer and surrounding tissue which was presumed healthy. The relatively small areas of each pathology, and discrepancies between the selections made by each of the two pathologists, limited the size and number of these manually-selected

regions. Plots of the average transmission spectrum from each of the three regions of interest show a similar overall spectral shape, differing predominantly in magnitude. The location of the ROIs and their average transmission spectra are shown in Figure 39.

The standard deviations between the three groups, shown in Figure 40, shows that the difference in magnitude is distributed fairly evenly across the full wavelength range, with different peak separation wavelengths occuring between different unhealthy tissue pathologies (hyperplasia, dysplasia and cancer) and between these and healthy tissue regions. There is most variation between the malignant stages between hyperplastic and cancerous tissue around 540, 785 and >890 nm though the higher wavelengths are less reliable being more prone to low SNR in this system due to depreciating hardware efficiency. There is most variation between the healthy and unhealthy tissue across the range 540 - 625 nm, potentially suggesting that residual haemoglobin, which has absorption peaks within this wavelength range, plays an important discriminatory role. Whilst the bit depth of the camera was 12-bit, the dynamic range of the data was notably lower than this suggesting that the aforementioned differences between the different pathologies may be significant.

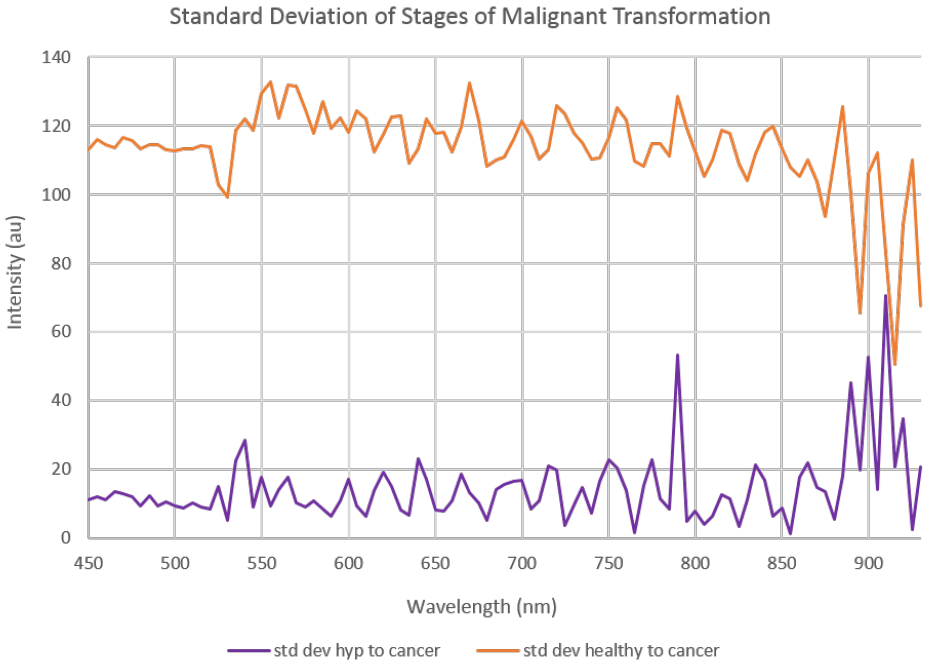

Figure 40: A graph of the standard deviations at each wavelength between HSI spectra collected from tissue classified by two histopathologists as hyperplastic, dysplastic or cancerous (purple) and healthy, hyperplastic, dysplastic or cancerous (orange) tissue from oral squamous cell carcinoma sample #8460.

As malignant transformation is often described as a progression from hyperplasia

to dysplasia to cancer, it may seem surprising that the hyperplasia and cancer spectra are more similar than the dysplasia and cancer spectra. However, that progression is in gradually worsening histological appearance, featuring increasing numbers of adverse feature accrual. In reality it is not necessarily a linear progesson from normal to cancer and new lesions often appear further along this traverse without apparently having transitioned throiugh each earlier descriptive step. The genetic errors accrued to change from normal to cancer can be acquired in any random order, only becoming significant when cellular function is disturbed and the disease presents clinically. Hence there is a huge potential range of DNA damage profiles that can present and develop into cancer from many and varied starting points. Additionally, alterations to the scattering spectra caused by different tissue characteristics cannot be separated, nor can they be assumed from context in diffuse scattering microscopy as the precise spatial position of individual scatterers cannot be recovered. While this setup approaches single scattering limits as it is carried out on thin tissue sections, it bears more similarity to diffuse scattering as the sandwiching of the tissue section between two calcium fluoride slides introduces multiple scattering events. This means that areas of hyperplasia with very high cell density, and therefore very high nuclear density, may report similarly to areas of low grade dysplasia with moderately elevated cell/nuclear density and nuclear size, for example, as an increase in the average size of scatterers will be produced in both cases. In extreme cases this could lead to the spectra of hyperplastic tissue appearing closer to neoplasia, than dysplastic tissue.

With such a broad range of malignancies, tissue types and numbers of fields of view (and associated errors) it is difficult to evaluate the efficacy of HSI for oral tissue segmentation from this study. It is clear that, in spite of strong background signal, unstained oral tissue can be separated from the background and from the OCT which surrounds it using this method, and that some internal tissue features are highlighted, such as keratin. It appears that hyperplastic, dysplastic and cancerous regions are also successfully separated from healthy oral tissue, though qualitative evidence from only one sample is insufficient to draw significant conclusions.

10.3 Limitations

10.3.1 Raman

Cluster analysis of the Raman data of these samples showed a distinct pattern of curved stripes, most likely arising from an Etalon effect. This is a potential risk of imaging older samples, and is a disadvantage of Raman relative to HSI (in addition to the large time investment Raman requires). As these were primary human samples and Raman requires specific mounting it was not possible to create new samples, and

despite efforts to change the coverslip spacing we were unable to eliminate the background artefacts.

The most disappointing outcome of the study was the discovery that many of the samples were degraded and produced unusable Raman data. Curved stripes covering the whole field of view dominated the k-means clustered maps, as shown in Figure 41. This meant that a three-way comparison between histopathology, HSI and Raman tissue maps was not possible. This means we cannot yet determine the compatibility of the HSI technique to be used in combination with Raman spectroscopy.

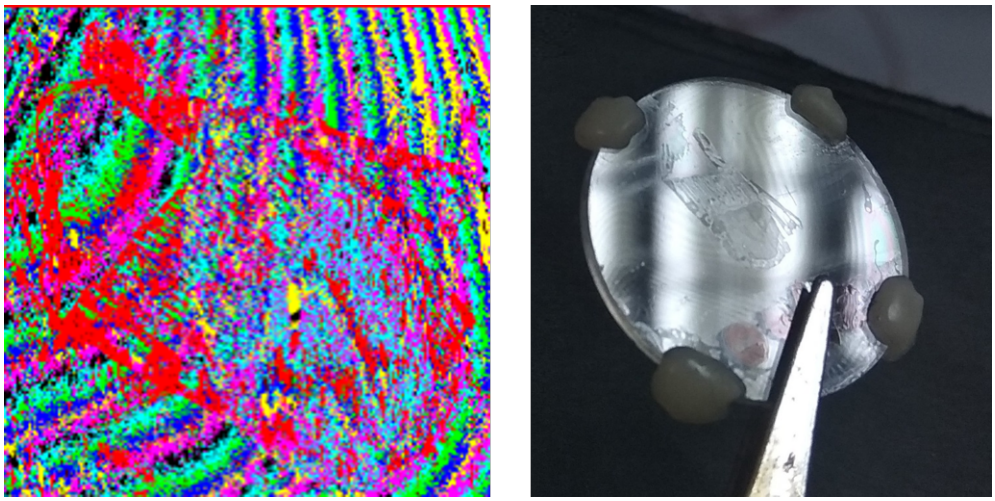

Figure 41: A k-means clustered Raman map of sample #4644, a tongue squamous cell carcinoma showing potential Etalon cavity background effects. (left) The effect is also visible by eye, demonstrated with sample #8460, an oral squamous cell carcinoma. (right)

Our best explanation is that multiple freeze/thaw cycles of the samples have slowly pushed the two calcium fluoride slides further apart, due to the unavoidable air entrapment between the two slides, creating something akin to an Etalon cavity which is typically made of a transparent plate with two reflective surfaces. Its transmission spectrum as a function of wavelength exhibits peaks of large transmission corresponding to resonance; a description which could also be ascribed to our Raman area scan maps. After attempts to heat and reseal the slides closer together failed, it was decided that new samples would be pursued for the next study.

Differences between the results from the remaining two techniques, histopathological cancer selection and k-means clustering of HSI stacks, come from three main sources:

1. Rotation/shear due to differences between imaging systems.

2. Physical distortion of different samples due to slice differences including damage from the delicate sectioning process.

3. High contrast background differences and the dominating effect this has on k-means clustering (also leading to tiling artefacts).

10.3.2 H&E

The poor quality of the RGB images of the H&E stained paraffinised tissue sections, especially the inhomogeneous illumination which led to patchy tiling effects, made it more difficult for the histopathologists to make accurate selections of regions of interest. A high magnification system with automatic tiling was used for the next study.

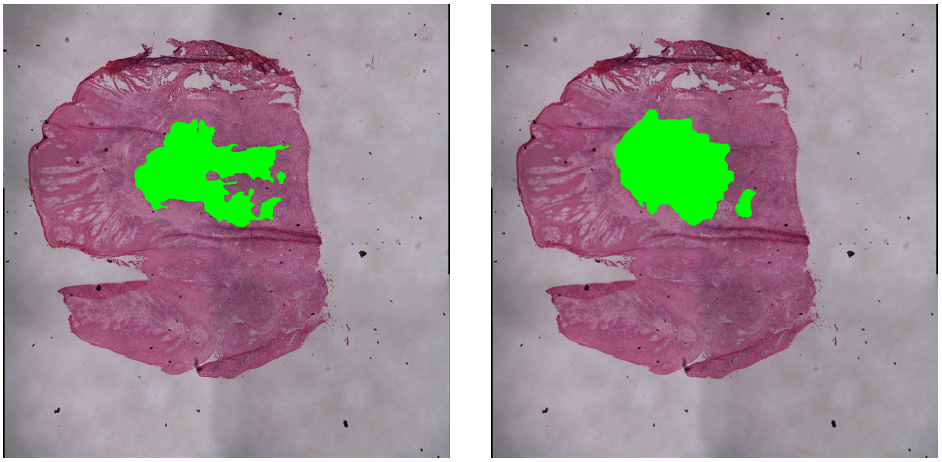

Figure 42: The RGB image of a H&E stained section of tissue containing oral squamous cell carcinoma, marked in green by two histopathologists, independently.

The tissue states hyperplasia, dysplasia and cancer, were demarcated quite differently by each histopathologist, as shown in Figure 42; an inevitable feature of the qualitative nature of pathology reporting.[133] This complicated the critical evaluation of the performance of the HSI and its segmentation. In order to quantify this gold standard qualitative diagnosis, one concensus diagnosis must be achieved across the panel of histopathologists for evaluation. This would also account for interobserver variation effects, which are substantial in practice. The regions of cancer selected by each histopathologist are shown in Figure 42. There is a core area of agreement to the centre and left of the areas selected. The areas of disagreement between the histopathologists cannot be taken as the gold standard diagnosis as, in each case, one histopathologist diagnosed the region as non-cancerous. For this reason, in the next study, it was decided that the concensus area only should be taken as the final cancer segment selected by this technique, and will serve as the ground truth diagnosis.

10.3.3 HSI

While the samples in this study (Results 2) were also partly imaged in the first study, it is unfortunately difficult to compare the datasets. In Results 1 (5.5), it was only possible to collect HSI of one field of view for each sample and, given that the complete datasets comprise up to 12 fields of view each and feature extraction is slightly different due to the changes in equiment between the two studies, it is very challenging to identify matching sample areas across the two datasets. We can, however, observe some general differences between the HSI data produced by the two setups.

The segment which provides the best internal tissue contrast in Results 1 has a spectrum which smoothly decreases with increasing wavelength, approximately exponentially. It was hypothesised that this resulted from Rayleigh scattering. The Halogen light source from Setup 1 (3) was replaced with a white light laser in this setup (Setup 2), with one primary motivation being extending the wavelength range towards the near infrared in order to maximise this scattering signal. Surprisingly, spectra with this shape are entirely absent from this dataset. The change from diverging sample illumination in the initial HSI system , to collimated illumination in this improved system inevitably decreased the acceptance of signal from higher scattering angles. This may explain the decreased sensitivity to the more isotropic Rayleigh scattering, relative to Mie scattering signal contributions which are forward-biased. This also suggests that a primary source of contrast in this study is absorption, rather than Rayleigh scattering. Spectral features are sharp rather than slow-changing which supports this assertion, though this may also be due to decreased overlap between spectral channels.[134] This improvement was generated by the filtering out of second order light from the monochromator and the increased spectral resolution afforded by the beam expander increasing the laser spot size on the monochromator grating. Small amounts of residual haemoglobin in vessels is the most likely source of absorption in unstained samples, though the samples appear colourless to the naked eye. There is often an increase in vasculature in the vicinity of tumours as the body responds to the tissue's increased demand for oxygen.[135] This is usually the primary source of contrast in similar work performed *in vivo*, or *ex vivo* on whole pieces of excised tissue, with blood.[136] If this were the case, we would expect troughs in the tissue spectra at 540 and 576 nm, and 555 nm around the absorption wavelengths of oxy- and deoxyhaemoglobin. A clear trough is present at approximately 530 nm, which could be a slightly shifted 540 nm oxyhaemolobin feature, though the absence of an associated absorption trough at 576 nm in the spectra of all tissue malignancy stages is surprising. While the plot of standard deviation between the healthy, hyperplastic, dysplastic and cancer spectra supports the assertion that haemoglobin

contributes key discriminatory absorption spectral features, its fairly flat trend across all wavelengths shows that all wavelengths contribute to the discrimination between the spectra of these different tissue pathologies, suggesting that diffuse scattering effects are also important. The broadly increased intensity of malignant tissue compared to healthy tissue suggests an increase in scattering affects in cancer, as anticipated.

However, knowing whether our primary source of contrast is scattering or absorption is not necessary in unsupervised segmentation such as k-means clustering, though it may help restrict the analysis parameters. Whether the Rayleigh scattering component is produced in the final dataset and isolated by its analysis, or not, we appear to retrieve relevant segmentation which is worthy of evaluation.

This apparently meaningful segmentation occurs in spite of three substantial system limitations; one relating to the equipment, one to the samples and one to the analysis software. Data taken with this setup (Setup 2) contained a large background component which was more complex and brighter than with the Halogen source used in Setup 1. A squircle pattern was produced which constantly evolved, with new squircles starting in the centre of the image and moving outwards concentrically. The possible reasons for this artefact are discussed in 9.2.1. Their position was not found to be consistent with any external input under our control, such as wavelength. This meant both that the illumination was not flat and that performing flat-field background correction would have depreciated the quality of the data.

The magnitude of these intensity fluctuations was almost a quarter of the magnitude of the feature fluctuations, decreasing the SNR of the tissue spectra and affecting the segmentation. It also meant that it was not possible to ensure the similarity of the background of different fields of view in order to tile them smoothly. Due to this, and perhaps due to the method chosen to calculate the regions where fields of view overlap, the tiling performed poorly and affected the data quality. The maximum value for each pixel was taken in overlapping tile areas. This was decided based on performance testing of a range of methods using an open source ImageJ macro written by Preibisch *et al.*[127] In our implentation, there was no removal of outlier values, as the unfamiliar data could not be thresholded without potential loss of useful information. This may have led to the introduction of a greater proportion of random noise in overlapping areas. The segmentation was affected because segments and their assigned colours do not match across field of view tile borders which introduces false boundary features into the data. This is relatively easy to overlook when visually interpreting the data but would be a significant source of error in a quantitative diagnostic approach.

Aside from effects introduced in the data collection and analysis, the samples themselves were challenging to coregister due to tissue cracking, folding, breaking and translating relative to other parts. This occurred both in the fresh frozen tissue sections

for HSI (and Raman spectroscopy), and the H&E stained, fixed tissue sections prepared for histopathological evaluation. Unfortunately, this is an unavoidable attribute of the delicate sample preparation process, but processing pipelines and software solutions must be implemented to minimise its impact on coregistration and segmentation quantification.

11 Conclusion

For identification of segmentation channels containing useful diagnostic information, and to evaluate the effectiveness of our HSI system compared to the current gold standard, a quantitative comparison must be facilitated between the HSI data and qualitative diagnosis from H&E stained sections by a histopathologist. The addition of a comparison between the HSI data and Raman spectra maps will show the compatibility of the two systems for a combination technique, though this was not unfortunately possible in this study due to the depreciatory effects of multiple freeze/thaw cycles on the samples and resulting Raman data. Automating comparison across imaging modalities is a complex problem. The primary difficulty is in coregistering the very different outputs from each of the three modalities, although tiling multiple fields of view making distinctive tissue edge features visible has made it possible to identify common tissue areas which was not possible in Part I (5.5). The coregistration process is further complicated by the high contrast and varying background introduced by the fibre-coupled white light laser which decreases the image quality and inhibits smooth tiling of multiple fields of view. Both equipment and software solutions were explored further to ensure this error was minimised in Part III. Similarly, areas containing sizeable sample defects must be excluded from the area of quantitative comparison. While poor coregistration, due to sample differences or the coregistration method itself, would negatively impact the outcome of the system evaluation, it would not affect the proposed combination HSI-Raman system itself in which only one sample would be utilised: No coregistration will be required once the efficacy of this technique has been established.

The outcomes of the preliminary data collection in Part I (5.5) necessitated three main changes to the setup: the addition of a white light laser (and fibre-coupling and collimation in place of an uncollimating liquid light guide) in order to increase SNR and the wavelength range, a beam expander to improve spectral resolution and a long pass filter to exclude second order light produced by higher order contributions from the monochromator at the longer wavelengths provided by the laser. The anticipated improvement in the precision of wavelength-specific spectral responses of the tissue, expected due to the removal of second order light and improved spectral resolution, was difficult to isolate as multiple alterations were made to the system simultaneously.

Nonetheless, the narrow bandwidth ("monochromatic") illumination was confirmed with a spectrometer and can only improve the quality of the hypercubes produced. Unfortunately, the successful extension of the wavelength range of the system did not obviously confer significant additional scattering contributions to the HSI data itself, given that Rayleigh scattering spectra were not identified and separated by the clustering software as they were in Part I.

The light was collimated in this setup by a reflective fibre collimator, whereas it was diverging directly from a liquid light guide in Setup 1. This may have resulted in increasing the detection of forward (Mie) scattered and unscattered light with Setup 2, whereas Setup 1 was more sensitive to more isotropically (Rayleigh) scattered light due to the different light path geometries through the sample to the camera. Nonetheless, the addition of longer wavelengths does add some useful information, as is clear from the approximately flat wavelength dependence of the discriminatory information between hyperplasia, dysplasia, cancer and healthy oral tissue shown in the standard deviation plot in Figure 39. The high frequency features present in the spectra also imply the anticipated improvement in spectral resolution and SNR conferred by the more powerful source and expanded beam incident on the monochromator grating.

This study has confirmed that the inclusion of tissue edges is worthwhile in improving the coregistration process and facilitating direct comparison between the RGB image and the HSI hypercube. This is expected to similarly improve the ease of comparing the HSI and Raman data which was excluded from this study due to damaged frozen samples. New samples were therefore collected to carry out a successful three-way comparison in Part III. Large variations between the three malignancies selected for this study, and their oral tissue types, made it more difficult to draw conclusions about the diagnostic value of the HSI segmentation. Constraining the samples to one common malignancy in one region of the oral cavity will allow spectral changes to be more specifically linked to the malignant transformation itself, on which diagnostic classifiers will be based in future work. This will also assist in constraining the analysis parameter options, which is vital when dealing with such large and information-dense datasets.

Part III

Economic Setup

White light lasers have a reputation for being temperamental, with the most common issues relating to damage to the microstructured optical fibres which are used to generate the nonlinear optical processes required to broaden the wavelength range spanned by the input pump beam source to create the supercontinuum. Additionally they are expensive, although their price has decreased rapidly over the past decade. For these reasons, including a white light laser in a setup intended for commercial production seemed ill-advised and when the nonlinear fibre burnt, the decision was made to try a cheaper and more reliable non-laser source; the Avantes Avalight-XE pulsed Xenon source. This substitute required several compromises, most notably decreased power, beam divergence, decreased spectral range and decreased spectral resolution due to decreased spot size on the monochromator grating and due to the entrance slit of the monochromator needing to be wider to ensure adequate power. Aside from the price and reliability, benefits included diminished interference patterns (improved further with the addition of a ground glass diffuser) which had been a major confounding element in the segmentation of the data and a simpler light path due to the source being fibre coupled which would also decrease the build costs for commercial production.

This fibre coupling also contributed to the reduction of field curvature. This increased the focussed area of each field of view so fewer were required for the imaging of a whole tissue section, speeding up the data collection and analysis processes; a vital consideration for a clinical tool. Imaging of full tissue sections, including distinctive tissue edges, in Part II facilitated comparison between HSI datasets and RGB images of H&E stained tissue sections for the first time, showing promising results. The development of a robust coregistration process was, therefore, a primary focus in Part III as part of the overall aim to quantify performance of the HSI system with respect to the gold standard.

The updated system was tested on HeLa cells with nuclear irregularities introduced by treatment with Paclitaxel. This informed the decision to also design texture analysis software to exploit the valuable spatial information present in the hypercube. This, and the previously outlined PCA denoising and k-means clustering software, were then quantitatively evaluated with respect to histopathology (and Raman spectroscopy) on healthy tongue and tongue squamous cell carcinoma thin tissue sections. This is the main study of the thesis and is the culmination of the system modifications and improvements detailed in Parts I and II.

Microplastics

Given recent concerns around bioaccumulation of particulate matter and microplastics and their potentially detrimental effects to health, including carcinogenic activity, preliminary work on their detection with HSI is included at the end of Part III (15.3). The COVID-19 pandemic this year brought single-use plastics into the spotlight, and additional fears surrounding the shedding of plastic fibres from surgical masks over time. The aim of this collaboration with Wright *et al* (Environmental Research Group, Imperial College London) is to develop a method to detect microplastics in tissue, with a view to screening human tissue biopsy samples for the presence of particulate matter and, especially, microplastics as an indication of exposure and uptake either via the lung or the gut. Particles < 2.5 μm in aerodynamic diameter may reach the deep airways where mucociliary clearance is no longer in effect, but they are commonly cleared by a variety of other biological mechanisms such as phagocytosis.[137] Larger particles which more commonly remain uncleared can lead to a variety of negative outcomes, such as inflammation.[138] The work of Wright *et al* has so far utilised Raman spectroscopy to image small sample areas, though the impracticality of this approach for a high-throughput, high resolution screening process is clear. A pre-screening system which locates potential particulate matter and microplastics first, and then collects Raman spectra from those same coordinates would be more suited to the desired application and aligns well with the overall aim of this work; development of a pre-screening process to enhance the clinical viability of Raman spectroscopy so its high specificity can be exploited. While environmental microplastics often contain dyes, colourless microplastics such as PMMA spheres are so devoid of clear spectral absorption features that they are often used to create tissue scattering phantoms. Their detection and subsequent segmentation in tissue with elastic scattering methods such as our HSI system was expected to be challenging.

12 HSI System And Software

12.1 Economic Setup

The light path adaptations that were made to to accommodate the fibre-coupled pulsed xenon source in the final, economic HSI setup are pictured below, in Figure 44.

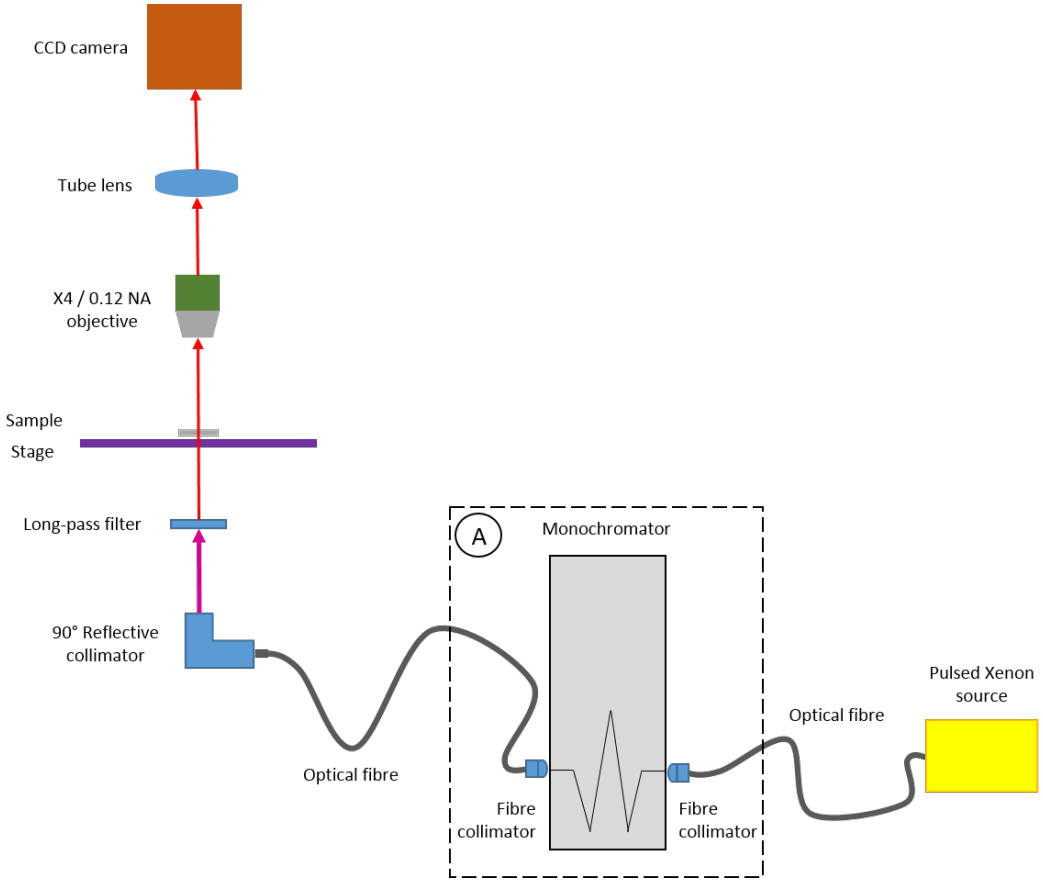

Figure 43: The light path of the final, economic HSI system with labelled area A corresponding with the photographs in Figure 44.

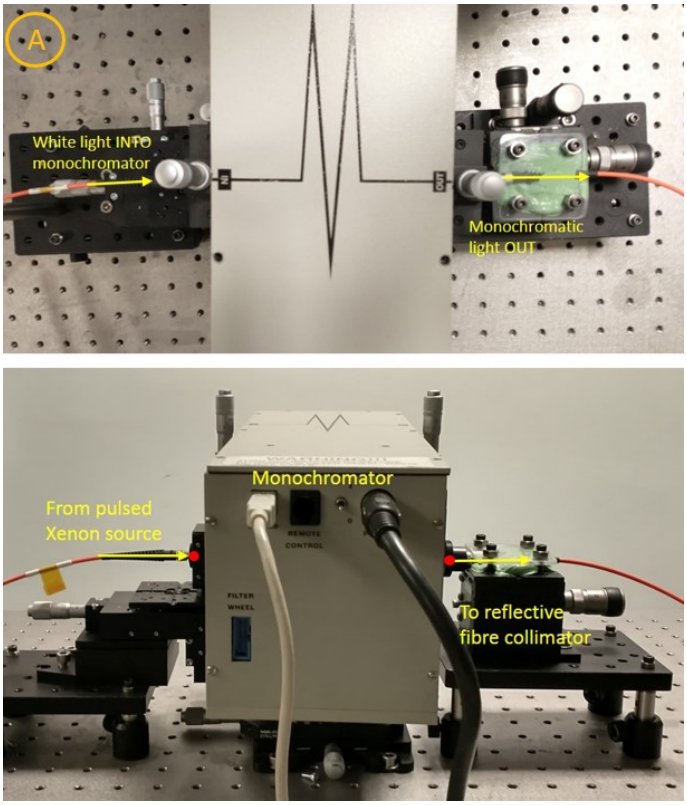

Figure 44: Top and side view of the light path adaptations to accommodate the new fibre-coupled pulsed xenon source. White light enters the monochromator through a fibre collimator (red dot in side view), monochromatic light is collected by a second fibre collimator (red dot in side view) into a fibre towards the 90° reflective fibre collimator which diverts the light up through the sample and to the camera. This figure corresponds with the section of the optical path labelled A in the optical path Figure 43.

The final setup is comprised of an Avantes Avalight-XE pulsed Xenon source triggered with an Avantes AvaSpec-2048-USB2 spectrometer, coupled with a Thorlabs 550 µm 0.22 NA step-index multimode fibre optic patch cable with a Thorlabs F220 SMA-A fixed focus collimator to a Newport Cornerstone 130 1/8m monochromator and holographic diffraction grating (1800 l/mm, 560 nm blaze, 320-1070 nm range). The monochromatic light from the output slit of the monochromator is collected with an 0.22 NA fibre collimator and 550 µm 0.22 NA single mode fibre pair from which the monochromatic light is directed via a second fibre collimator into a fibre, up to the sample stage by a Thorlabs 90° reflective fibre collimator through a ground glass diffuser (1500 Grit. Thorlabs), through a 785 nm long pass filter (Semrock) in a motorised filter flip mount (MFF101, Thorlabs), then collected by a x4, 0.13 NA Olympus objective held in the turret of a Zeiss Axiovert microscope and focused onto the CCD of a ORCA-Flash4.0 V3 Digital CMOS camera by an achromatic doublet lens (f=100 mm, Thorlabs, AC508-100-A-ML, Ø2", ARC: 400-700 nm).

The Avantes Avalight-XE pulsed Xenon source required at least 30 minutes to stabilise after switching on, before use. Whilst this could appear problematic for the proposed application and the high speed system it necessitates, in a professional laboratory this system would be rarely turned off, negating any significant time losses to equipment start-up.

12.2 Final Data Collection Software

The data collection software remained unchanged from Part I and facilitated the collection of a full hypercube with user-specified wavelength inclusion, autoexposure, filter flipper control for exclusion of second order light from the monochromator and concurrent dark field stack collection. As in Part II, collimated sample illumination meant that the use of pre-defined focal positions at each wavelength was no longer required.

For full details, refer to 3.2.

12.3 Final Segmentation Software

12.3.1 PCA and K-Means Segmentation

The home-built Matlab PCA dimensionality reduction and k-means clustering software outlined in 7.3 was utilised again for this study.

As literature suggest that classifiers which use both spectral and spatial information perform best, a spatial analysis technique was also employed for testing.1.2.2

12.3.2 Texture Analysis

Texture analysis characterises image regions by quantifying the "textures" of which it is comprised. As most samples will contain areas of different textures, locating and demarcating the boundaries between these textures can provide useful segmentation. Physical textures such as smoothness or bumpiness are identified by looking at the spatial variation of pixel intensity values.

A statistical method for texture examination which has proven value for tissue segmentation is the grey-level co-occurrence matrix (GLCM). According to the original definition of the GLCM, made by Haralick *et al*, for an image I with N greyscale levels, is a square matrix G of order N where the (i,j)th element of G represents the number of instances of a pixel with value i being adjacent to a pixel with value j. Adjacent in this case means "neighbouring" and can be in any of the four possible directions relative to the pixel in question.

This adjacency criterion can be altered to search for larger texture features by increasing the magnitude of the offset from 1 (neighbouring pixels) to n (pixels which are n pixel "steps" away from the pixel in question).

When a sample is known, or contains well-defined textures, the offset direction and size can be tailored to the sample. For instance, stripes of known size and separation in Type I collagen fibres could be targeted specifically. In exploratory work such as this, however, it is common to create GLCMs for multiple offset directions and sizes (or a "patch") and take the average for each texture feature calculated for each different offset. The most commonly used texture features are the Haralick textures: contrast, correlation, energy and homogeneity which are defined below.

Texture Feature	Definition
Contrast	Magnitude of local greyscale variations
Correlation	Probability of the occurrence of the pixel pair in question
Homogeneity	Uniformity of pixel greyscale values
Energy	Closeness of the distribution of elements in the GLCM to the GLCM diagonal

Table 3: Descriptive definitions of the four Haralick textures contrast, correlation, homogeneity and energy.

A programme was developed in Matlab to create a texture map of each wavelength slice of a HSI stack. For each wavelength slice, the four texture values are calculated for a patch of a fixed size, P, which is moved over the full image with fixed step sizes in x and y, S_x and S_y. Each texture value is converted into a grayscale pixel intensity value which is assigned to area of size S_x x S_y the in the top left of the patch before the patch is moved (by S_x or S_y) to its next calculation position.[9] This windowing approach enhanced the effective spatial resolution of the technique without making the patch too small to detect textures, which often require a degree of repetition for recognition. The resulting output of this texture analysis was four texture image stacks, each the same size as the input hypercube but with a spatial resolution in x and y of S_x and S_y , respectively.

These four texture image stacks were then segmented by k-means clustering, but with no dimensionality reduction as all the tissue components were required to explain a high percentage of the data variance.

[9]This top left patch filling necessitated an additional translation to coregister other modalities before tile scoring. Filling in the centre of the patch would have prevented this and will be implemented in future.

13 Raman System and Software

13.1 Raman Setup

The Renishaw InVia Raman spectrometer (Renishaw, Gloucestershire, UK) was used to collect Raman area maps of the tissue with Renishaw's newer WiRE 4.4 software. Renishaw have increased the speed of high SNR area scan acquisitions by utilising a cylindrical lens to distribute a high power laser line across the sample surface. This technique and software accommodations, known together as Streamline™, allows several point spectra along the laser line to be collected simultaneously without depreciating spatial resolution in x or y and decreases the total time required to scan the laser across the full surface of the sample.[128]

13.2 Analysis Software

The home-built Matlab PCA dimensionality reduction and k-means clustering software outlined in 7.3 was utilised again for this study, though Raman does not require field of view tiling as it is a point scanning technique.

14 Materials and Methods

14.1 HeLa Cells with Paclitaxel

A preliminary investigation into whether the final HSI system could differentiate between cancer-like and healthy cells was undertaken with paclitaxel-treated and untreated HeLa cells. Paclitaxel is a chemotherapeutic drug used to treat several cancers by inhibiting the proliferation of cancer cells by inhibiting the mitotic spindle. This means that it can change healthy cells in ways which make them appear cancer-like: At low concentrations (5 - 10 nM) it causes aberrant mitosis, with ~ 55% of cells being multinucleated after 20h of treatment, and higher concentrations cause terminal mitotic arrest with aneuploidy until cell death.[139, 140] Aneuploidy (cells with an abnormal number of chromosomes) is a feature of most cancers and can be detected by elastic scattering spectroscopy as elastic scattering results from changes in refractive index including changes caused by nuclear chromatin content.[141, 142]

14.1.1 Sample Sourcing and Preparation

(Samples prepared by Mark-Alexander Gorey)

1. Human cervical cancer cells (HeLa) were thawed and cultured in Dulbecco's Modified Eagle Media (DMEM) 1% Penicillin and Streptomycin (PS, Sigma[1]P0781), 10% Foetal Bovine Serum (FBS, HyClone-SV30160.03) and 1% L-Glutamine (L-Glu, Sigma[1]59202C).

2. Cells were transferred to a T25 flask (Cellstar-690175) and incubated at 37°C and 7% Carbon dioxide.

3. Passaging cells were trypsinized with 10% trypsin (Sigma-T4174) diluted in Phosphate-Buffered Saline (PBS, Sigma-D8537), for 5 min once cells reached 80% confluence.

4. Complete DMEM was added to neutralize trypsin and centrifuged at 1200 rpm for 3 min.

5. The supernatant was aspirated, and cells were resuspended in 5ml of complete DMEM.

6. The cells were then seeded on two glass microscope slides and kept in separate dishes, submerged in DMEM until they formed a 70% confluence monolayer.

7. After two days, the medium was replaced with

   (a) Fresh, paclitaxel-free complete medium in one dish.

   (b) Complete medium containing paclitaxel (Sigma Aldrich Paclitaxel, Taxus sp. - CAS 33069-62-4) at a 20nM concentration in the other dish.

8. After 6 hours at 37 ° C the paclitaxel-containing medium was replaced with fresh, paclitaxel-free medium. (Lower concentrations of paclitaxel applied for longer time periods, as suggested for aberrant mitosis, resulted in cell death and inability for the treated HeLa cells to approach confluence.)

9. Both the healthy and paclitaxel-treated HeLa cells received an additional change of medium (paclitaxel-free).

10. Incubation was continued until 100% confluence was achieved for both samples. (In practice, the treated cells did not exceed 80% confluence.)

11. Both samples were then fixed with PFA 4% for 15 min.

12. They were then mounted with Merck Millipore FluorSave Reagent (345789) and topped with a glass cover slip.(If possible, the samples would have been left unfixed but our small stage meant that keeping them submerged in media was not possible, and the imaging protocol was too long for them to be left dry.)

14.1.2 Data Collection

Hyperspectral Imaging

Hyperspectral image stacks were collected with a x4, 0.13 NA objective lens (Olympus) across a wavelength range of 521 - 899 nm in increments of 6 nm producing a 64 slice hyperstack for each field of view (FOV). Optimal camera exposure time for each wavelength were read in from a pre-collected text file throughout the hyperstack collection process. At 785 nm, the long pass filter was inserted into the optical path to filter out second order light. Collection of corresponding dark field stacks was interleaved with raw image collection by triggering the automated shutter of the monochromator and taking an unilluminated, but otherwise identical, image at each wavelength to capture and correct for background light, dead camera pixels and other dark-visible imaging artefacts. Additionally, one flat field stack was collected per objective lens, to allow for manual completion of the flat field background correction protocol. This resulted in a total, unoptimised data collection run time of approximately 12 minutes per repeat. Unfortunately the camera rolling shutter, which we were unable to change with this VB.NET software, in combination with the new pulsed source required 6 repeats of each dataset in order to correct for the horizontal striping artefacts. This unfortunately increased total data collection time per hypercube to over an hour, though this challenge is not insurmountable for future permutations of the system.

For full data collection instructions and reference to the software GUI, please refer back to 4.3.1.

Data Processing

The data was then processed by the method outlined in 9.2.1. The home-built shift and brightness fine correction ImageJ macros (9.2.1, 9.2.1) were utilised as previously described. The change of illumination source from a laser to a pulsed-Xenon source removed the constantly varying squircles which prevented the full background correction of the data in Part II so the full flat-field background correction procedure could be used as intended (9.2.1). The dark field corrected stack was divided, slice-by-slice, by the (dark field corrected) flat field stack producing a hypercube ready for analysis. No field of view stitching was required as only one was collected.

14.2 Tongue Squamous Cell Carcinoma and Healthy Tongue Samples

14.2.1 Sample Sourcing

Patient tissue samples and data were provided by Guy's & St Thomas' Head & Neck Biobank – part of the KHP Cancer Biobank, which is supported by the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award and Guy's & St Thomas' NHS Foundation Trust. The biobank has ethical approval from the East of England - Cambridge East Research Ethics Committee (18/EE/0025).

This study utilised six healthy tongue sections and six tongue sections which carried a diagnosis of squamous cell carcinoma, selected from the sample set of Dr. Manickavasagam. The reasoning behind enforcing these two criteria is discussed in section 11. Imaging was undertaken "semi-blind" in that I was aware of the diagnosis for each tissue section (in order to select appropriate samples) but not what features on the slide led to the diagnosis.

14.2.2 Sample Preparation

Two subsequent (or as close as possible) tissue sections were cut from each of the listed tissue sections. One 4 µm section in frozen section fixative was stained with H&E on a glass microscope slide and topped with a glass coverslip, for histopathological review, and one 10 µm frozen section was sandwiched between two round Raman grade calcium fluoride slides for hyperspectral imaing and Raman spectroscopy.

For full protocols, please refer back to 4.2.2.

14.2.3 Data Collection

Hyperspectral Imaging

Hyperspectral image stacks were collected with a x4, 0.13 NA objective lens (Olympus) across a wavelength range of 521 - 899 nm in increments of 6 nm producing a 64 slice hyperstack for each field of view (FOV). Optimal camera exposure time for each wavelength were read in from a pre-collected text file throughout the hyperstack collection process. At 785 nm, the long pass filter was inserted into the optical path to filter out second order light. Collection of corresponding dark field stacks was interleaved with raw image collection by triggering the automated shutter of the monochromator and taking an unilluminated, but otherwise identical, image at each wavelength to capture and correct for background light, dead camera pixels and other dark-visible imaging artefacts. Additionally, one flat field stack was collected per objective lens, to

allow for manual completion of the flat field background correction protocol. This resulted in a total, unoptimised data collection run time of approximately 12 minutes per repeat. Unfortunately the camera rolling shutter, which we were unable to change with this VB.NET software, in combination with the new pulsed source required 6 repeats of each dataset in order to correct for the horizontal striping artefacts. This unfortunately increased total data collection time per hypercube to over an hour, though this challenge is not insurmountable for future permutations of the system.

For full data collection instructions and reference to the software GUI, please refer back to 4.3.1.

Data Processing

The data was then processed by the method outlined in 9.2.1. The home-built shift 9.2.1 and brightness fine correction 9.2.1 ImageJ macros were utilised as previously described. The change of illumination source from a laser to a pulsed-Xenon source removed the constantly varying squircles which prevented the full background correction of the data in Part II so the full flat-field background correction procedure could be used as intended (9.2.1). Additionally, flatter focus across the field of view increased its useable area and decreased the number of fields of view required to image a full sample clearly, though the stitching process in the analysis software was still governed by global coordinates retrieved using the ImageJ field of view stitching macro, run after data collection as described in 9.2.1.[127]

Raman Spectroscopy

The Raman data collection was performed in the same way and using the same parameters as in Part II; a full tissue area scan taken in Streamline™mode with the WiRE 4.4 software using a x20 (0.45 NA) objective, a 785 nm diode laser at 100% power and a 2s exposure time, a 600 l/mm grating over a ~ 981 – 1719 /cm wavenumber range with a spectral resolution of 1.9 /cm per pixel and a Ybin parameter of 5 yielding effective spatial resolution of 50 µm in the x-direction and 13.7 µm in the y-direction.

For full data collection instructions and parameter rationales, please refer back to 9.2.2.

Histopathological Images

Full slides were scanned at x40 magnification using a Hamamatsu Nanozoomer 2.0HT with a LX2000 Lightning Exciter source.

14.2.4 Assessment of Tissue

Overview of Comparative Study

The improvements to the system, software and comparison methods culminatede in the development of an analysis workflow which is automated at almost every stage and requires little human input bar the necessary histopathological diagnoses. This will allow the performance of the system to be evaluated in a simple and repeatable manner after future alterations and could be adapted for release, allowing its use by other groups with different systems. It uses ImageJ and Matlab only, both widely used in the field and with open source options available.

As the development of this system was piecemeal in response to the outcomes of the smaller studies in Parts I and II, and the three datasets (HSI image stacks, Raman spectral area maps and RGB histopathology images) require different treatment, its workflow may be unclear at this stage. For clarity, a flowchart depicting the procedure from data collection to quantitative evaluation is shown below (Figure 45). Some of the steps, such as the affine transformation for coregistration, will be introduced for the first time later in this subsection (14.2.4). The reader is encouraged to refer back to this flowchart throughout Part III, whenever necessary.

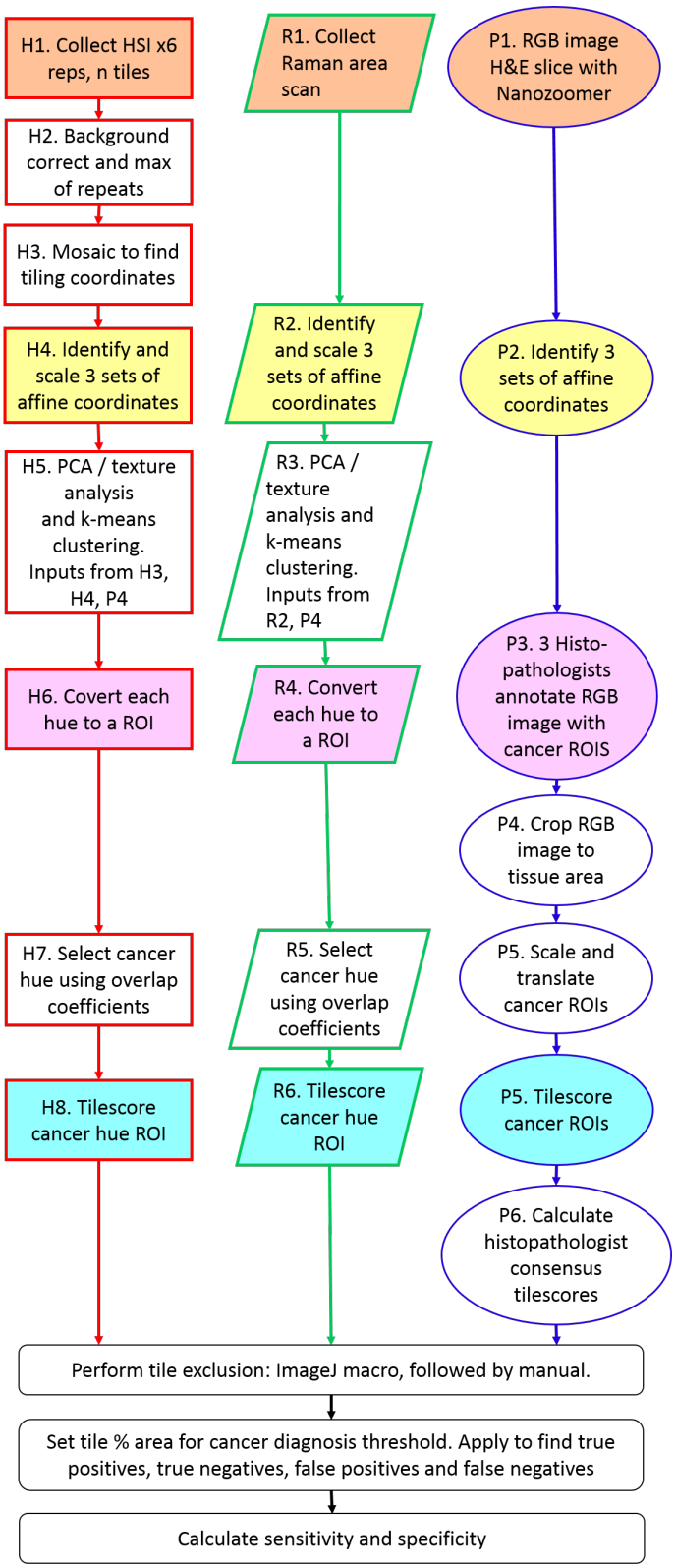

Figure 45: A workflow outlining the process from collection of HSI (red rectangle), Raman (green parallelogram) and RGB image (blue oval) collection, through their correction, segmentation and performance analysis. Colour filled boxes indicate where the data from the three modes is in the same format or undergoes the same process.

Cross-Modal Coregistration

The affine transformation-based coregistration software processes are outlined and depicted below (Figures 46 and 47), separately to the segmentation software as they are performed at different stages on each dataset. The HSI stacks are transformed to match the H&E tissue area before analysis as the texture analysis "pixels" were designed to align perfectly with the scoring tiles, thus providing an immediate diagnosis for each texture tile according to whether or not it is filled with the cancer texture hue. The Raman data was segmented before affine transformation as it was prone to transformation errors due to its low spatial resolution (discussed further in 15.2.4) and propagating these errors throughout the segmentation would have lead to more time consuming re-analysis. Additionally, scaling up the size of the image int the affine transformation to match the larger, higher resolution modes inevitably leads to some scaling errors which were better handled by running the segmentation before performing the affine transformation.

Affine Transformation

In order to relate the manually-segmented H&E stained colour images, the segmentation of the hyperspectral image stacks and the Raman maps, the tissue areas must be coregistered. This allows for pixel-wise comparison by correcting for differences in the field of view, translational offsets and rotational differences inherent in imaging a sample on a circular slide (calcium fluoride) on different equipment, along with any transformations that take place within the analysis software. The most commonly used multimodal coregistration technique is affine transformation; a linear mapping technique used to correct for geometric deformations and distortions such as cropping, resizing, rotating, shear mapping and translating. Points, straight lines and planes are all preserved.

In order to calculate the required affine transformation in Matlab, three sets of well-spread feature coordinates which were identifiable in each modality were manually located in GIMP using the same overlaying method as in Part II (9.4). The manual identification of fiducial markers can be quite time consuming and is vulnerable to human error. Unfortunately, given the numerous complications (circular calcium fluoride slides, sectioning artefacts, changes in tissue morphology across non-subsequent slices etc.), human involvement for this step was unavoidable. This non-automated step, however, does not impact on the automation of the hyperspectral diagnostic workflow, only on this exploratory comparative work. The coordinates of these three points were input in the home-built Matlab software and used to calculate the affine transformation to coregister the HSI stacks and the Raman spectroscopy area scans of the unstained tissue section to the RGB image of the corresponding H&E stained tissue section. The

affine transformation settings were set such that the transformed images would also be padded in x and y to match the RGB image of the H&E section. These settings are vital to automating the coregistration of the images after they have been transformed and their implementation in Matlab is included below.

$$Aslice = imwarp(slice, tform, 'OutputView', imref2d(size(P)));$$

where "slice" is the slice of the hypercube or Raman spectral stack which will be affine transformed and padded, "tform" is the previously calculated affine transformation, "Aslice" is the affine transformed and padded image (HSI or Raman), and "P" is the RGB image of the H&E stained section being used as a reference for the post-affine padding.

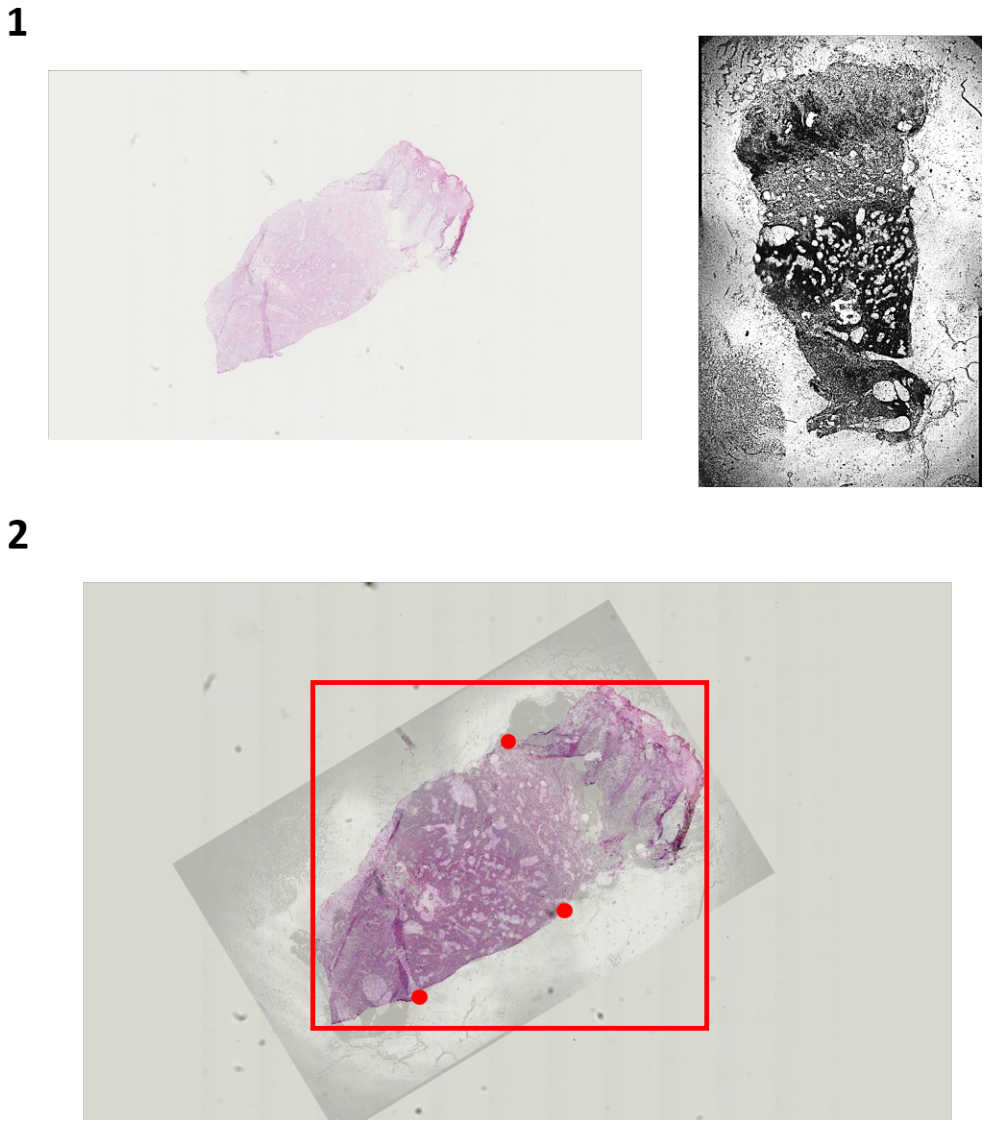

Figure 46: 1. Separate H&E and HSI images. 2. Coregistered by selecting 3 common reference markers in each image and performing an affine transformation in Matlab.

This affine transform was then applied, slice by slice, to the HSI and Raman data. The H&E image was chosen as the reference image as the histopathologists' cancerous region selections on the stained sample represent the gold standard diagnostic procedure to which all other approaches should be compared. Additionally, the cancer regions of interest selected in ImageJ by the histopathologists are difficult to transform, and subsequently analyse, without loss of important information.

Tissue Edge Cropping Macro

Due to the large amount of white space surrounding the tissue in in all modes after padding to match the RGB image of the H&E stained section, the (now matching) datasets were cropped to avoid time-consuming processing of empty tiles. An ImageJ

macro was developed which generated an automatic tissue outline by converting the image from type RGB to HSB (hue-saturation-brightness) and performing an auto-thresholding of either the hue or saturation images (this inconsistency was a result of staining variability between different samples). This selection was converted into an ImageJ ROI (region of interest), from which the size and position of its bounding rectangle can be extracted using the macro language function "getSelectionBounds", and saved as a text file. This macro was run on all of the RGB histopathology images before coregistration software was run on the other two modes so that the bounding rectangle details could be input into the coregistration software for cropping of the affine transformed and padded HSI and Raman datasets.

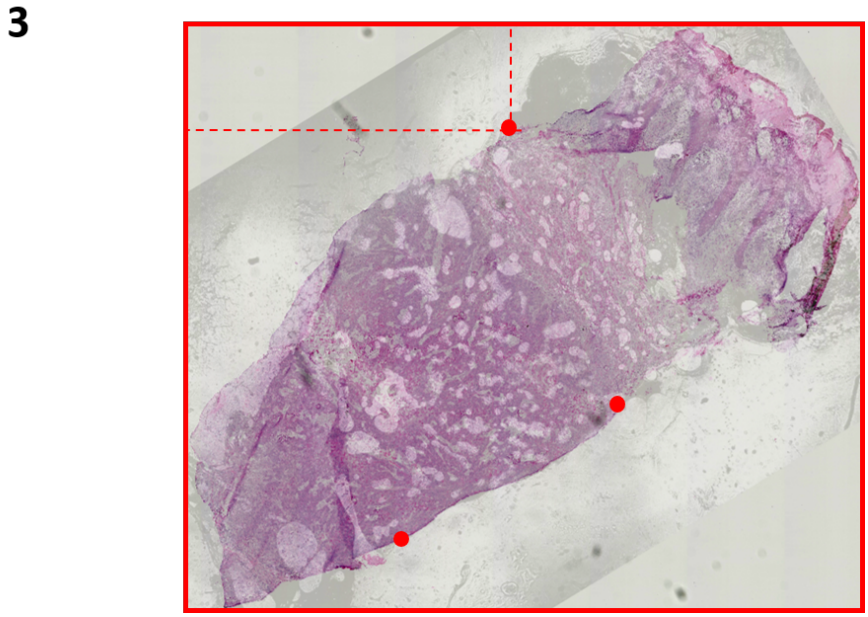

Figure 47: 3. The H&E image was cropped to the outer boundaries of the tissue with an ImageJ macro. By recalculating the positions of the common affine transform reference coordinates with reference to the new origin (top left corner of the cropped H&E image) the affine transformed HSI image was cropped to match.

This step required a little human input as the bounding rectangle details were collected in ImageJ and had to be manually transferred to Matlab for use in the coregistration software, however this is a feature of the evaluation system which would not be present in the proposed combination HSI-Raman benchtop diagnostic system. In such a system, the HSI and Raman datasets would be co-occurring by nature as they would be collected in succession on the same tissue section on the same stage.

Segmentation

Segmentation by Histopathologists

Cancerous regions were selected by histopathologists according to the protocol detailed in 9.3.1, but for this study three histopathologists independently made these demarcations. This allowed interobserver variation to be accounted for.

Interobserver variation is an indicator of one of the greatest downfalls of the current histopathological process; it's lack of repeatability. However, as histopathology is the current gold standard for the diagnosis of oral cancer, the performance of new techniques must be measured relative to this. To try and overcome the potential impact of interobserver variation on the accuracy of this comparative study, the consensus diagnosis was calculated from the regions of interest selected independently by the three histopathologists using method detailed in 14.2.5.

Segmentation of HSI Data

Texture Analysis

Texture analysis was performed as an additional technique to attempt to enhance segmentation of the HSI data and the software is outlined in 12.3.2. As recommended in exploratory work, a GLCM was created for multiple offset directions (0°, 45°, 90° and 135°) and sizes (1, 2, 3 and 4 pixels) for a patch size of 396 pixels which equates to 798 µm. The mean texture value across all offsets was then taken for each texture feature (contrast, correlation, homogeneity, energy) and the top left 99 pixels (200 µm) of the patch (the texture "pixel") were then filled in with a grayscale value equal to the offset-mean patch texture value. The patch was then moved across the HSI hypercube slice with a step size of 99 x 99 pixels (200 x 200 µm) and this process was repeated across the full HSI image stack to produce four texture image stacks per HSI dataset.

Patch sizes which are too large will not be sensitive to early cancer invasion and patch sizes which are too small may not detect patterns or textures. With these considerations in mind, parameters were utimately chosen so that each complete texture "pixel" matched the position and dimension of each scoring tile, the dimensions of which were selected with reference to the different patterns of cancer invasion and the average size of squamous epithelial cells, discussed in full in 14.2.5.

PCA and K-Means Segmentation

As in Part II (9.3) the component number for the PCA denoising software was selected was based on the turning point of the cumulative plot of percentage data variance described by each component, shown in Figure 48. The processed HSI (processing detailed in 14.2.3) was denoised with PCA, however the four texture stacks produced from

the processed hypercubes were not, as a clear noise threshold for exclusion could not be determined from the graph of variance explained by each component which exhibited no clear turning point.

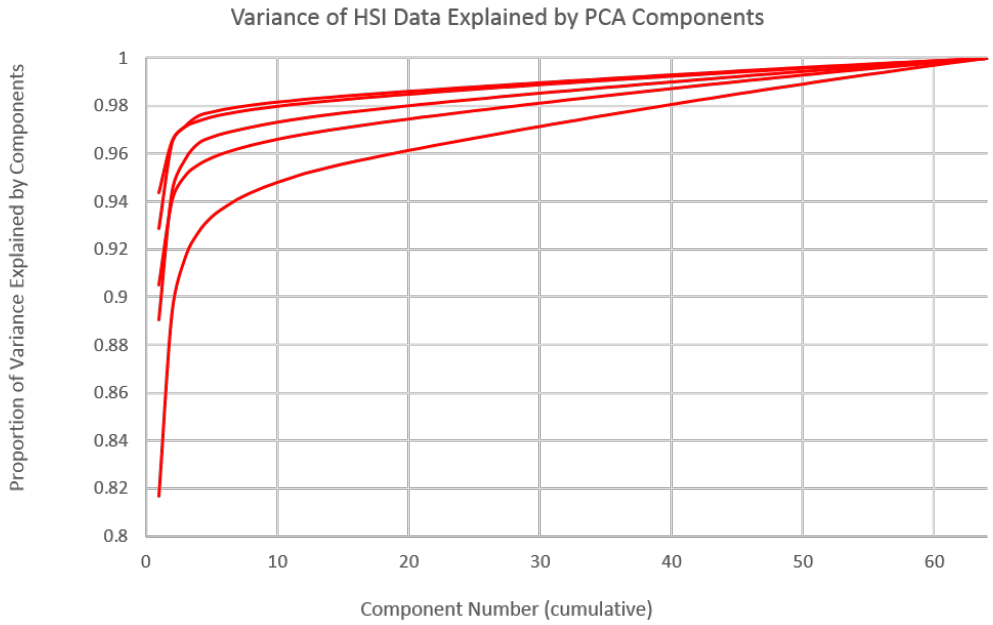

Figure 48: A graph of the variance proportion of the HSI data of the six tongue squamous cell carcinoma samples explained by each component as determined by the Matlab software detailed in 7.3.2.

A component number of 5 was selected for the processed HSI accordingly: Smaller than the component number of 7 selected in Part II, this suggests either than that this data has a lower SNR resulting in fewer valuable components, that use of the white light laser in Part II added additional information which was not available with the pulsed-Xenon source or that the two additional components in the results from Part II correspond with intense noise features such as the varying background squircles which could not be successfully removed.

The processed HSI and the four texture stacks produced from those hypercubes were finally segmented by k-means clustering. The cluster number for each sample was assigned according to the assumption of one background cluster per field of view tile and six main tissue components found in cancerous oral tissue sections, as discussed in 7.3.3.

Segmentation of Raman Data

The Raman data was noise filtered by PCA with a component number which explains 95% of its variance: The high SNR achieved with Raman scattering means it requires minimal noise filtering and this 95% threshold is commonly used.[143] It was

then k-means clustered with a fixed cluster number of 7, dictated by the rule of thumb successfully applied in previous work; 6 tissue components and one background contribution per field of view (of which there is always 1 for the Raman data).

Automated Cancer Region Selection from HSI, HSI Texture and Raman K-Means Segmentation

Converting K-Means Segments to Regions of Interest

The single affine transformed, cropped grayscale output image produced by the coregistration and segmentation software for each mode (PCA dimensionality-reduced HSI, each of the four HSI textures, PCA dimensionality-reduced Raman) needed to be quantitatively compared with the gold standard consensus segmentation afforded by the histopathologists in the format of ImageJ regions of interest (ROIs). Basic manipulation of these ROIs is well provided for in ImageJ but exporting them for use in other software is difficult. Therefore an ImageJ macro was written for converting the segmented grayscale images to tissue segment ROIs. Before running, a sample area of each background segment must be selected by the user; this is used to reject background ROIs and automate selection of the tissue area. At this stage, the macro identifies every grayscale value within the tissue boundaries and, using thresholding, separates areas of each grayscale value into a separate segment ROI.

Samples which this protocol performed worst on were those with little tissue contrast, poor tiling performance, high contrast background features or artefacts close to the tissue itself. For example, presumed Etalon cavity effects which were detrimental to the Raman datasets in Part II (10.3.1) also affected sample #4644 causing an intense, striped background across the scan area which was visible over the tissue area in some places. These stripes were assigned to two clusters which were also present within the tissue area in places, inhibiting the macro from being able to detect the tissue edges clearly: one of these two clusters is shown in orange in Figure 49

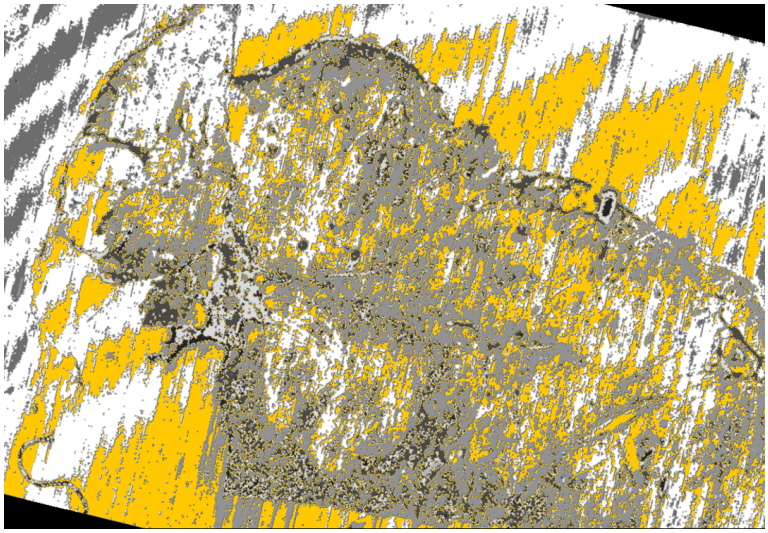

Figure 49: Etalon cavity effects in the Raman area scan of tongue squamous cell carcinoma sample #4644 inhibiting automated tissue area selection. One segment, found in the background and the tissue area, is selected in orange.

In these few cases, some manual intervention was necessary to tailor the macro to overcome these artefacts.

Selection of Cancer Region

As there are limited examples in published literature of hyperspectral scattering microscopy on thin, unstained oral tissue sections the aim of this work was, in part, to identify diagnostically relevant features of the data. Without this prior knowlede of the diagnostic features of the spatial and spectral facets of this data, segmentation based on HSI diagnostic classifiers was not possible. It was therefore decided that the data would be arbitrarily segmented using k-means clustering and, for each sample, the most similar segment to all three of the histopathologist-selected cancer regions would be identified using a home-built ImageJ macro based on overlap coefficients. The overlap coefficient (OC) is a measure of similarity between two sets (A and B) and is defined as follows.

$$OC(A,B) = \frac{|A \cap B|}{min(|A|,|B|)}$$

In this case, A would be the gold standard histopathologist selected cancer ROI and B would be an individual segment ROI of the grayscale segmented image (k-means clustered HSI, HSI texture or Raman). (The histopathologist consensus cancer ROI and the intersection of these two ROIs can both be found by using the "AND" command in the ROI Manager and the smallest ROI can be determined by comparing their measured areas.) For each such image, the overlap coefficient was calculated for each segment

ROI and the segment with the highest similarity to the gold standard was selected as the "cancer ROI".

Fundamentally, k-means clustering separates areas with the greatest spectral differences into individual segments and therefore, retrospectively identifying the spectral features which delineate each segment from the others is an acceptable way to begin the process of identifying potential diagnostic classifiers.

14.2.5 Quantitative Cross-Comparison

Tile Scoring

The affine and cropping step facilitates direct comparison between corresponding tissue areas of different sections of the same sample. From this point, each coregistered modality for the same sample was broken into identical tiles of a fixed size, as shown in Figure 50, and the percentage coverage of each tile by its cancer ROI (tile score) was calculated. An ImageJ macro was developed to perform this process which we refer to as "tile scoring".

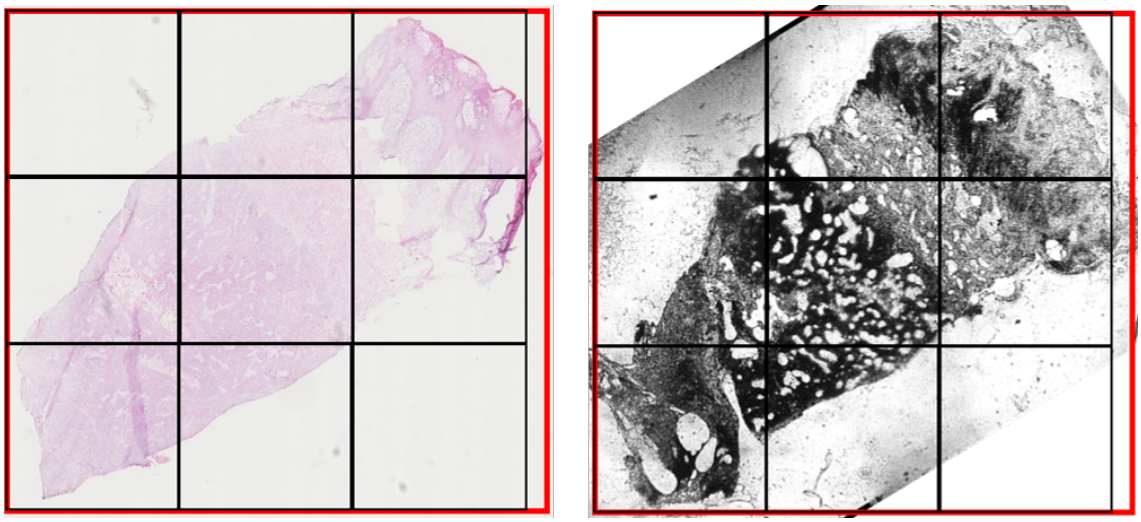

Figure 50: A coarse tiling representation to show how the affine and cropping facilitates easy tile-wise comparison between the different modalities. Red box indicates image edges and small black squares represent tiles.

To decide the tile size for quantification of the cancer segments, we referred back to our original aim; the early, automated detection of oral cancer for improved survial rates. While just one epithelial cell the wrong side of the basement membrane would still be classified as infiltrating cancer, the likelihood of that single cell being successfully captured in a biopsied tissue section and detected by the histopathologist (our gold standard reference diagnosis) is low. (Additionally, this same tile size must be assigned to the step size, or effective pixel size, of the texture analysis. A single squamous cell

can range in size from ~ 40 - 60 µm and, with a spatial resolution of ~2µm, a single cell patch size is unlikely to return reliable texture values.) Therefore we looked to cancer infiltration patterns for a more practical suggestion. Once the basement membrane is breached and cells start to stream from the epithelial to the mesenchymal side, which is the definition of infiltration, this stream of cells can infiltrate by single-cell or collective migration, classified as less, or more, than 15 cells across, respectively.[9] Very small tile sizes would lead to less reliable cross-comparison due to decreased likelihood of identical areas being contained within matching tile numbers across different imaging modalities due to the human errors which are inherent in the coregistration process. With this in mind, it was decided that collective infiltration would be considered in this study. Therefore a patch size of 200 µm was selected; equating to a cluster of 16 squamous cells (arranged 4 x 4) with an average size of 50 µm. This tile size was fixed and tile scoring was only performed on tiles completely within the tissue area.

Histopathologist Consensus Calculation

At the tile scoring stage, the tile becomes an effective pixel with one descriptive value, its score, and we no longer take into consideration the positions of the cancer regions within each tile. Therefore for each tile, the minimum percentage coverage assigned by any of the three histopathologists was taken as the consensus diagnosis.

Tile Exclusion

One of the main limitations of the investigation has its roots in the affine transformation step. As, in practice, there are between 10 and 50 µm (1 - 5 tissue sections) between the unstained and corresponding stained tissue sections, the morphological differences can be pronounced. Other sectioning artefacts which we encountered included folding, wrinkling and cracking, with or without additional translation or rotation of the separated area. Affine transforms are unable to account for any of these artefacts. Therefore, a two-stage process of excluding tiles falling within areas of these artefacts was developed in ImageJ.

Automatic

The tissue area ROIs created previously for the H&E stained sections and the unstained sections used for the HSI and Raman data collection (in the tissue edge cropping macro (14.2.4) and the segment ROI creation macro (14.2.4), respectively) were saved and re-used in this stage to identify large areas of disagreement. The included tiles were initially restricted only to those fully contained by the union of the tissue outline ROIs from all three modes. The numbers of included tiles were identified in an ImageJ macro and returned as a list which was used in combination with Excel filtering and

sorting functions to include only the the tiles fully within the tissue area boundaries in all modes, rejecting fragments and dislocations and partially accounting for changes in tissue shape between sections..

Manual

Manual tile exclusion was not intended from the outset, but the samples required it. While large areas of dislocation were effectively excluded by the automated tile exclusion macro, some artefacts such as tissue folds did not meet the auto-exclusion criteria and were manually identified and excluded before implementing the included tile list by filtering and sorting in Excel.

Evaluation of Tile Scores

The performance of the tissue segmentation by HSI and Raman was evaluated in the form of sensitivity and specificity of the technique. Sensitivity records the true positive rate: the proportion of cancerous tissue which is correctly identified as cancerous. Specificity returns the true negative rate: the proportion of non-cancerous tissue which is correctly identified as non-cancerous.

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

where TP (true positives) are tiles where the cancer segment from the mode under evaluation and the histopathology cancer segment both cover more than DT% of the tile area, TN (true negatives) are tiles where the cancer segment from the mode under evaluation or the histopathology cancer segment both cover less than DT% of the tile area, FP is where the cancer segment from the mode under evaluation covers more than DT% of the tile area but the histopathology cancer segment does not and FN is where the cancer segment from the mode under evaluation covers less than DT% of the tile area but the histopathology cancer segment covers more than DT% where DT is the defined diagnostic threshold (discussed and selected in 15.2.4).

14.3 Microplastics

14.3.1 Sample Sourcing

Particulate Matter from Air Filters

The particulate matter sample was collected from an air quality monitoring supersite at Marylebone Road, over a 24 hour sampling period.[144] Collection used a Thermoscientific Partisol 2025i with a PM10 size selective inlet which selects particulate matter (PM) below 10 D µm in aerodynamic diameter which is generally considered to be of "inhalable" size.[145]

Microplastics in Mouse Pulmonary Tissue

Mouse pulmonary tissue sections collected from lungs dissected from terminated control mice were borrowed in collaboration with Wright *et al* (Environmental Research Group, Imperial College London). Pulmonary tissue was selected as the respiratory system, along with the digestive system, are hypothesised to be the main sites for microplastics (and particulate matter) to enter, and potentially cause damage in, the body.

14.3.2 Sample Preparation

Air Filter Samples

The PM present in the air sample was added to water (concentration was not specified as dropcasts do not preserve concentration) and a 10 µl drop was dropcast onto a calcium fluoride slide (Crystran CaF_2 20mm diameter, 0.5mm thickness, polished Raman grade). The sample was left to dry slowly (without heat) to avoid dramatic coffee ring effects, and on an orbital shaker (Stuart) at a low speed to avoid particle aggregation which would erroneously enhance detection of small particles.

Microplastic-Perfused Mouse Pulmonary Tissue

1. Lungs were dissected from terminated control mice. Frozen tissue (A) was inflated with OCT and formalin-fixed tissue (B) was inflated with PBS.

2. Plastic microbead mix was made up to 0.01% final concentration in inflating media. It contained equal concentrations of polymethylmecrylate (PMMA) beads (5 – 27 µm), polyethylene (PE) beads (10 – 27 µm) and polystyrene (PS) beads (4 and 10 µm).

3. Frozen tissue (A) was inflated with OCT:PBS (1:1) containing plastic microbeads (as detailed above) and then frozen in OCT using dry ice.

Tissue was then processed as follows:

1. Frozen samples were cut into 10 µm sections at -20 °C using a cryostat (Bright, UK).

2. Sections were affixed onto Superfrost Plus slides (Thermo Scientific Menzel, UK) and left to dry overnight in the fume cabinet.

14.3.3 Data Collection

Hyperspectral Imaging

Hyperspectral image stacks were collected with a x4, 0.13 NA objective lens (Olympus) across a wavelength range of 521 - 899 nm in increments of 6 nm producing a 64 slice hyperstack for each field of view (FOV). Optimal camera exposure times for each wavelength were read in from a pre-collected text file throughout the hyperstack collection process. At 785 nm, the long pass filter was inserted into the optical path to filter out second order light. Collection of corresponding dark field stacks was interleaved with raw image collection by triggering the automated shutter of the monochromator and taking an unilluminated, but otherwise identical, image at each wavelength to capture and correct for background light, dead camera pixels and other dark-visible imaging artefacts. Additionally, one flat field stack was collected per objective lens, to allow for manual completion of the flat field background correction protocol. This resulted in a total, unoptimised data collection run time of approximately 12 minutes per repeat. Unfortunately the camera rolling shutter, which we were unable to change with this VB.NET software, in combination with the new pulsed source required 6 repeats of each dataset in order to correct for the horizontal striping artefacts. This unfortunately increased total data collection time per hypercube to over an hour, though this challenge is not insurmountable for future permutations of the system.

For full data collection instructions and reference to the software GUI, please refer back to 4.3.1.

Data Processing

The data was then processed by the method outlined in 9.2.1. The home-built shift and brightness fine correction ImageJ macros (9.2.1, 9.2.1) were utilised as previously described. The change of illumination source from a laser to a pulsed-Xenon source removed the constantly varying squircles which prevented the full background correction of the data in Part II so the full flat-field background correction procedure could be, and was, used as intended (9.2.1). No field of view stitching was required as only one was collected.

15 Results and Discussion

15.1 Preliminary Study - HeLa Cells Treated with Paclitaxel

In order to study the differences between the healthy and paclitaxel-treated HeLa cell HSI data, they were grown in monolayers across separate glass microscope slides, fixed and topped with a cover slip, and imaged side-by-side in the same field of view to ensure intensity variations across different datasets did not inhibit direct comparison. As paclitaxel induces mitotic arrest, the cells did not form an even monolayer after treatment; this is an important limitation of this preliminary study as cell packing impacts elastic scattering spectra.[146] This, in combination with the criterion that only cells near the slide edges could be included in one field of view, impacted the unbiased selection of representative areas of cells from each sample.

Therefore, the centre strip was selected from each sample (in one dataset), positioned only to avoid out of focus areas at the edge of the field of view and areas at the slide edges with very poor cell coverage. As previously, segmentation map colours are randomly assigned to each cluster within each dataset, however their assignation can often be inferred from the image and associated sample knowledge. The (light) blue segment appears to demarcate cell nuclei which are evidently different after treatment with paclitaxel (Figure 52). Therefore, masks of the blue (nucleus) segment were created from both of the sample areas described previously, to facilitate collection of nuclear area and perimeter statistics. These masks for paclitaxel-treated and untreated Hela cells are shown as binary maps in Figure 51, to the right of their associated segmentation maps.

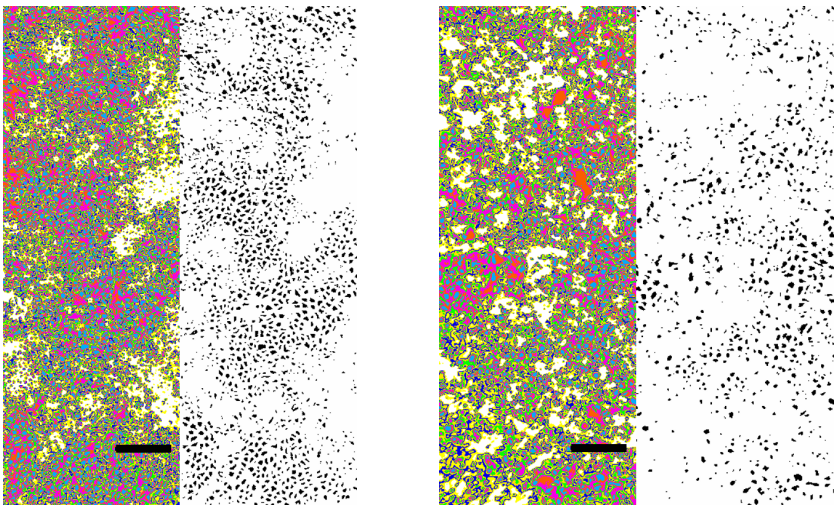

Figure 51: K-means cluster map with corresponding mask of nuclei of of healthy HeLa cells (left) and paclitaxel-treated HeLa cells. Scale bar represents 150 µm.

Analysis of nuclear area and nuclear perimeter were peformed on these whole regions, in spite of varied confluence of the paclitaxel-treated cells. The paclitaxel-treated nucleus data is slightly less reliable than the healthy cell data due to a lower cell count in this region due to the aforementioned inhibited proliferation.

Feature	Healthy	Paclitaxel-Treated
Mean Nuclear Area (px)	9.9	12.0
Nuclear Area Std. Dev. (px)	8.5	11.1
Mean Nuclear Perimeter (px)	11.3	11.7
Nuclear Perimeter Std. Dev. (px)	6.6	7.15

Table 4: Mean values of nuclear area and perimeter in healthy and paclitaxel-treated HeLa cells, and the standard deviations of these feature measurements for each group (1 d.p.).

Both the average nuclear area and nuclear perimeter, and their standard deviations, show an increase with paclitaxel treatment. This supports the hypothesis that our HSI system with k-means clustering is able to detect differences between healthy and cancerous cells, as paclitaxel treatment mimics several of the cellular changes of cancer such as aneuploidy. This is also in agreement with studies comparing nuclear features in oral squamous cell carcinoma, our target sample set, and healthy oral tissue.[147] Smaller areas containing successful monolayers of healthy untreated, and paclitaxel-treated HeLa cells were isolated to visually demonstrate the nuclear area and perimeter differences identified.

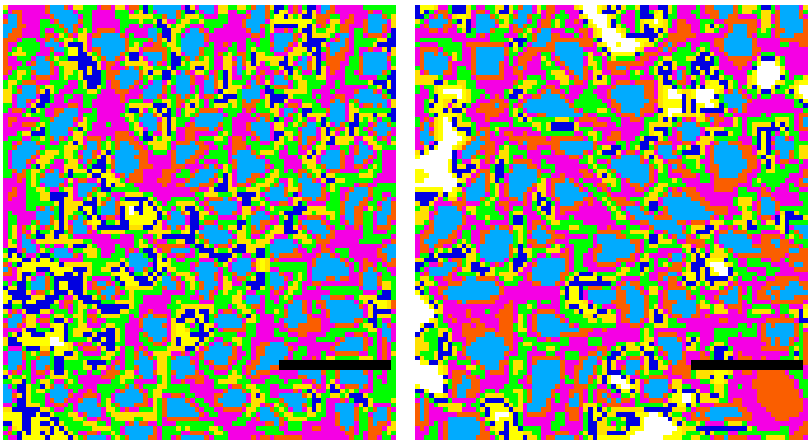

Figure 52: K-means cluster map of an area of healthy HeLa cells (left) and HeLa cells that have been treated with paclitaxel. Scale bar represents 50 µm.

The cell packing irregularity in paclitaxel-treated cells is also evident in this figure (Figure 52), suggesting that spatial analysis will be a useful addition to the analysis process.

Unfortunately, the inhibited proliferation of the paclitaxel-treated HeLa cells led to patchy cell coverage and varying cell packing densities which inhibited broader segmentation of the field of view to show global differences between the two different cell states. If cell coverage had been even, showing larger areas similar to Figure 52, statistics regarding segment coverage across the whole slide could have been utilised to more coarsely segment the image. For example, patches or tiles with more than 40% covered by the blue hue (assumed nuclei and suggesting a high nuclear-to-cytoplasmic ratio which could be assumed cancerous/cancer-like) could be selected automatically. Fortunately, cell confluence is not a concern when utilising tissue sections.

This preliminary work evidencing that differences between healthy and cancer-like cells can be detected by this HSI system when segmented by k-means clustering provided hope that this system would perform equally well, or better, on tissue.

15.2 Main Study - Tongue Squamous Cell Carcinoma Segmentation

Due to the huge variety of tissue types and structures within the oral cavity, variation between the dissimilar components of healthy samples could be as large as variation between healthy and diseased (e.g. malignant) tissues. For this reason, a subset of malignant oral tissue samples has been selected which are expressing the most common serious oral malignancy, squamous cell carcinoma, and are from the region of the oral cavity most commonly affected; the tongue. These malignant areas were compared with healthy regions contained within the same samples, and also with healthy tongue samples. Limiting the source of variation to the cancer itself allows us to demonstrate the capability of the system to quickly delineate between healthy and malignant or unusual tissue.

Performance was predominantly evaluated using receiver operator curves (ROC) which plot true positive rate against false positive rate. They are a popular way to represent the diagnostic ability of a system which makes a binary classification; in this case a cancerous or non-cancerous tile of tissue.[148, 149] ROC curves illustrate the compromise between sensitivity and specificity by plotting how they alter with the application of different discrimination thresholds; in this case, what percentage coverage of a tile by a cancer segment classifies the whole tile as cancerous. (The segmentation of all modes produce tiles with partial cancer coverage, except for the texture analysed HSI data which classifies full tiles due to application-informed patch size constraints.) Selecting a diagnostic threshold requires a compromise between false positives and false negatives, each with different negative clinical implications. The former assigns a cancer diagnosis to a healthy sample, triggering unnecessary and expensive diagnosis

and treatment advancement steps. The latter has an even graver outcome, missing the cancer and labelling the tissue as healthy, diverting the patient away from the required treatment path and decreasing their chance of survival. The practical and moral implications of selecting a diagnostic threshold are substantial and the selection should be made only when the diagnostic technique has been fully optimised and found worthy of clinical translation. Potential appropriate diagnostic thresholds are briefly discussed later, nonetheless. In order to evaluate the performance of the technique without ascribing a fixed diagnostic threshold, the area under the receiver operator curve, or AUROC, score represents the separability performance of the technique across all possible discrimination thresholds. The AUROC score can take any value between 0 and 1, and the higher its value, the more successfully the two classes (in this case cancer and non-cancer) are separated. Usually a classification model or system with an AUROC score > 0.60 is considered to have fair class separation capacity, >0.70 is good, >0.80 is very good and >0.90 is excellent and 1 is a perfect. Using this performance metric allowed the system to be fully evaluated without having to select an appropriate diagnostic threshold.

15.2.1 Interobserver Variation

Many investigations into potential histopathology adjuncts or replacements reference high interobserver variation but do not account for it in practice. This may be due to fear of depreciated diagnostic performance of systems compared to a more conservative gold standard, in spite of the potential for increased accuracy and precision.

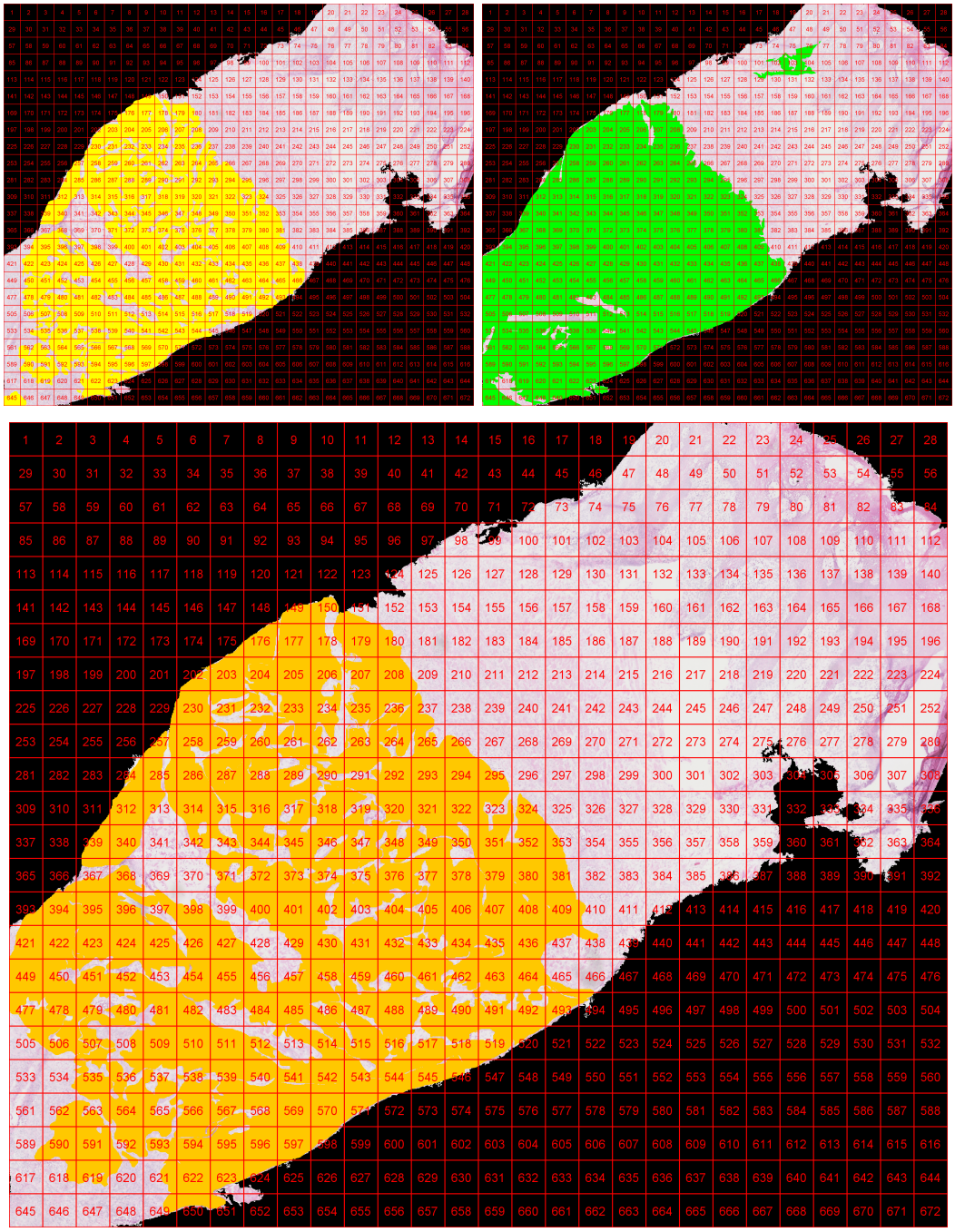

Figure 53: RGB image of tongue squamous cell carcinoma sample #9806 with demarcation by each of the three histopathologists with scoring tiles overlayed in red and numbered for reference (one image enlarged for clarity of tile numbers). Note partial tile coverage by cancer, requiring a discrimination threshold for assigning tile-wise labels of true positive, true negative, false positive and false negative for the other modes, in comparison with this gold standard. Each tile is 200 x 200 µm in size.

Looking at the cancer segments selected by the three histopathologists for sample #9806 in Figure 53, there are areas of substantial disagreement. The rectangle of unselected area in the bottom left of the green-selected image, suggests that some of

those areas of disagreement could be erroneous rather than truly representatative of a different diagnostic opinion. While the ROI selection technique (9.3.1) was reasonably simple and robust, it was vulnerable to some errors. Predominantly, human errors in precise and accurate selection of intricate cancer infiltration patterns, but also a feature of ImageJ ROI selection which automatically closes open ROI edges with a straight line once the mouse is released. This is likely to be the cause of the rectangular region across tiles 506 - 511 of the green-selected image, and may have overwritten some small areas which were intentionally not selected in this same sample. This contributed to the decision to utilise the histopathologist consensus diagnosis, rather than the mean or maximum which would contain increasingly more potentially erroneous contributions. Additionally, the consensus diagnosis is often used as the gold standard in studies which look into differences in the diagnoses provided by multiple histopathologists.[150] Even though the yellow and orange-selected cancer regions appear very similar for this sample at first glance, there are differences at closer inspection. For example, in the bottom left column of the dataset (tiles 421, 449, 477, 505, 533, 561, 589, 617 and 645) show large difference with the yellow-marking histopathologist diagnosing cancer only in the very bottom tile, the orange-marking histopathologist diagnosing cancer only in the top half of that column, and the green-marking histopathologist detecting almost complete cancer coverage across these tiles.

Different varieties of Kappa values, especially Cohen's Kappa, are often cited as measures of interobserver agreement.[151] Unfortunately Cohen's Kappa metric are only suitable for two-way comparisons within discrete data. Percentage agreement between histopathologists is another popular measure of interobserver variation, occasionally favoured in diagnostic imaging as it provides more clinically relevant (absolute, rather than relative) information, in spite of not accounting for chance where Kappa statistics do.[150] It unfortunately requires that potential outcomes are binary which necessitates the setting of a diagnostic threshold. Rather than make underinformed assumptions about this threshold, the intra-sample and inter-sample mean standard deviations of the percentage coverage of cancer selected by each of the three histopathologists was evaluated.

The mean intra-sample standard deviation ranged from 2.4% in tongue squamous cell carcinoma sample #9890 to 12.6% in sample #4644. The smaller standard deviations often correlated with samples containing less cancer. Where there is more healthy tissue in a sample containing cancer, the number of zero-cancer-coverage tiles increases. These are often the simplest to locate and the least contentious between different histopathologists. Tiles containing cancer, or near a main tumour island, are likely to require more complex discriminatory decisions by the evaluator such as dif-

ferentiating between cancer and dysplasia or hyperplasia. Additionally, the manual selection of the cancerous region is subject to human error whereas fully healthy tissue tiles are not. It is important to remember that this ImageJ selection process (and its associated human error) is not representative of the diagnostic process undertaken by histopathologists on a daily basis, but is instead a necessary feature of a comparative study such as this.

Taking the mean of these intra-sample standard deviations gives us an estimate of how varied the percent cancer coverage is, per tile, across all of the cancer-containing samples: This value was evaluated as 6.7%. While comparison with other studies of interobserver variation would be informative, this is difficult due to the range of metrics utilised. The effect of interobserver variation is not negligible in this study, which is in general agreement with literature and supports using the consensus area of cancer coverage per tile as the gold standard.

15.2.2 Sources of Error

When we evaluate performance of an imaging technique by calculating sensitivity and specificity we are, in reality, evaluating the imaging technique (within certain parameters dictated by the equipment and other practical limitations, such as the magnification used), the process by which the resulting images are quantified (including the choice of analysis parameters) and the process by which these quantitative results are compared to the reference mode. Performance metrics can therefore be substantially impacted by errors in any of those steps. From the flowchart depicting the process from data collection to final evaluation (Figure 45), we can identify the following steps which may depreciate the apparent performance of the HSI screening technique:

1. Processing the multiple HSI repeats to correct for the rolling shutter (potential loss of useful information by taking the maximum at each pixel across all repeats)

2. Tiling multiple HSI fields of view (potential acceptance of more noisy pixels, raising the SNR, by taking the maximum at each pixel in tile overlap areas)

3. Finding common coordinates between the three datasets (HSI, Raman and RGB image of H&E) for affine transformation to allow cross-comparison.

    (a) Sampling errors and sample artefacts will also become apparent in this step, as even a perfect coregistration technique cannot match and align data from two different samples (HSI and Raman data on the fresh frozen tissue section and the RGB image of the H&E stained, paraffin-embedded, separate tissue slice, each with their own artefacts).

4. Analysis parameters chosen based on small bodies of evidence and/or logical assumptions, such as k-means cluster number and spatial texture offset step sizes and directions.

The table below shows which of these limitations has particularly affected each sample, referenced using the numbers from the list above, and links them to a specific sample feature. This will be explored further in the sample-wise evaluation of the HSI-segmentation technique.

Sample Number	Dominant Limitations
9806	3a - Broken tissue
2217	3, 3a - A sample artefact caused by a cracked slide
4644	3a - Minimal healthy tissue present in sample
2577	2 - Tiling artefact. 3, 3a - Fragmented frozen section
9890	2 - Four fields of view so increased tiling artefact
4734	2 - Four fields of view so increased tiling artefact. 3a - Many folds in H&E section

Table 5: A table outlining the limitations (from the numbered list above) most affecting the HSI datasets of each malignant sample.

The magnitude of each contribution to the total error is difficult to estimate, though we may exploit the fact that the HSI and Raman spectroscopy were performed on the same sample to gain some additional insight. When we evaluate the performance (sensitivity and specificity) of the HSI with respect to the gold standard histopathology, the outcome is likely to contain contributions from all of the error sources outlined above. However, if we evaluate the performance of the HSI with respect to the Raman "cancer" segment (selected by highest overlap coefficient with respect to the histopathologists' combined cancer ROI), it will contain fewer errors due to sample artefacts as both techniques are performed on the same tissue section, though errors may still have been introduced in the affine transformation used to coregister the two modalities. Unfortunately, these performance metrics would only reflect agreement between the HSI and Raman cancer segmentation. However, the higher the sensitivity and specificity of the Raman data with respect to the histopathology (indicating accuracy with respect to the gold standard), the more diagnostically relevant this would be. As Raman spectroscopy has proven diagnostic power, with cancer diagnostic sensitivities and specificities of over 90% encountered regularly in literature, we may consider this a valid comparison.[152]

15.2.3 Raman Performance Evaluation

Our Raman spectroscopy system, in combination with the process by which we compared it with the gold standard histopathology-based diagnosis, performed quite poorly relative to leading literature garnering an AUROC score of 0.69, where scores of 0.95 have been reported with the use of a similar segmentation technique and a PCA-hLDA

based diagnostic classifier.[152] In this study, exposure time was sacrificed in order to utilise the broad wavenumber range required to include the most discriminatory spectral features. The application our system was designed for necessitated the use of full tissue sections. Raman spectroscopy is a notoriously slow technique, and imaging tissue sections of ~1 cm^2 with an adequate pixel size, exposure time, spectral range and resolution can take over 24 hours. This decision to compromise on exposure time may have been more detrimental given the number of freeze/thaw cycles our samples had experienced prior to final data collection, which can lead to diminished intensity of chemical peaks (requiring maximising of the available signal such as by increasing the exposure time).[120] However, given the clinical application this system is designed for, even with the Raman scan area decreased substantially by a HSI pre-screening and segmentation process, speed must be considered at every stage. All chemically specific sample mapping modalities which do not require markers, such as Raman and FTIR spectroscopy, are comparably slow by nature. While developments in laser technology are beginning to enhance the speed of FTIR, at a cost, and its resulting appeal, Raman spectroscopy facilitates the most convenient sample preparation for work on full tissue sections and is therefore the most suitable option. Therefore, practicality imposed a 16 hour scan time limit resulting in a two second limit on the exposure times used for all samples, and therefore depreciating the SNR and potentially impacting the segmentation performance. In the proposed HSI-Raman combination system, the impact of time-consuming nature of Raman spectroscopy will be significantly mitigated by the sparser sampling facilitated by the HSI pre-screening process. Therefore this limited, two second exposure time will not be a limiting factor in the final combined system and its effect on the data is a feature of this preliminary study only.

Nonetheless, random sampling from the spectra of cancerous regions of our samples show features at many of the characteristic wavelengths suggesting that the data is accurate, if not particularly high in quality. Spectral differences between healthy oral tissue and oral squamous cell carcinoma peaks were most commonly noted at 1341 cm^{-1} and 1265 cm^{-1}. The former has been correlated with dead cells, which has clear histopathological relevance.[153] In spite of angiogenesis in the tumour vicinity to accommodate its rapid metabolism, necrosis occurs in tumour centres as access to nutrition and oxygen is restricted as the tumour grows. The peak indicates the presence of adenine and guanine of DNA; adenine has been proven to be a reliable biomarker of oral cancer discriminant in Raman spectroscopy by Dai *et al.*[154] The second peak, at 1265 cm^{-1}, is associated with protein. Increased protein and DNA in cancer cells aligns well with the global picture of cancer pathology.[54]

15.2.4 HSI PCA and K-means Segmentation

Performance Evaluation

The ROC curve in Figure 54 shows the power of HSI imaging with PCA-based dimensionality reduction and k-means clustering as a standalone tongue squamous cell carcinoma discriminatory technique as it compares it only to the gold standard histopathologist-segmented RGB images of the corresponding H&E stained sections (ground truth).

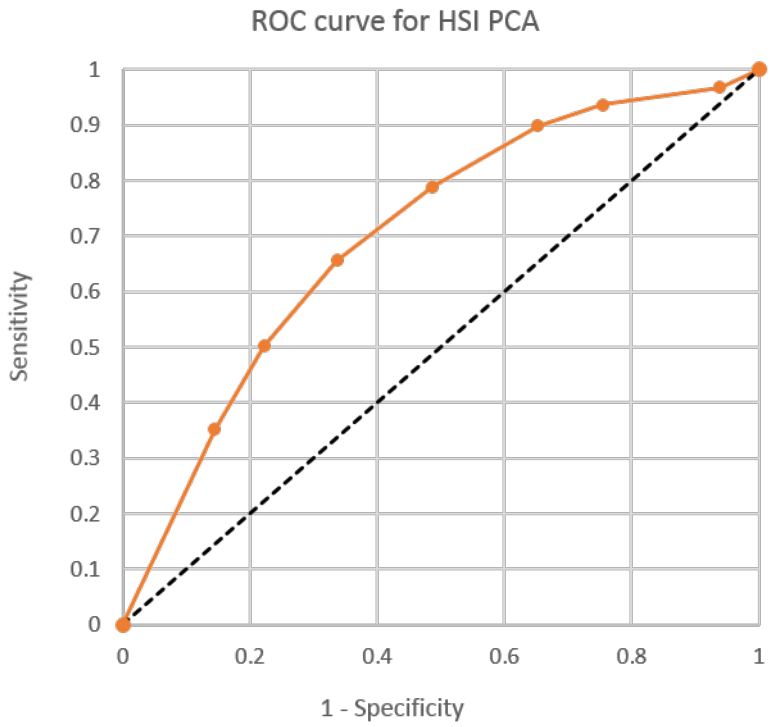

Figure 54: ROC curve for segmentation by k-means clustering of dimensionality-reduced HSI of tongue SCC tissue with respect to the consensus histopathologist segmentation of the subsequent tissue section. Black dashed line represents a worthless diagnostic test (chance).

The AUROC score is 0.70 which means that k-means clustering of dimensionality-reduced HSI has good capacity for separation of tongue squamous cell carcinoma from healthy tongue tissue. Given that the HSI and Raman spectroscopy both use the same tissue section, we would expect better performance of the HSI with respect to the Raman rather than the histopathology. The segmentation produced by the HSI system compared to the Raman clustering as the ground truth did garner a slightly higher AUROC score of 0.73, though the improvement is not as marked as anticipated. This can be interpreted in three potential ways, either the coregistration between those two modes has performed badly or the Raman-based segmentation is of poor quality, or it is

is incompatible with the HSI-based segmentation suggesting that their discriminatory spectral features come from dissimilar tissue areas or components. The Raman datasets are more prone to coregistration errors as the raw datasets are very small (in x and y) due to the low spatial resolution and must, therefore be scaled up in size during the affine transformation. This means that any small human errors introduced in the affine coordinate selection process are also scaled up, resulting in worse tissue area matching. This is likely a major source of error, along with the poor quality Raman data discussed previously (15.2.3), in the reporting of performance of HSI with respect to Raman segmentation and Raman segmentation with respect to histopathology as the ground truth.

While comparing the HSI-histopathology performance in tandem with the HSI-Raman performance can illuminate the important sources of error in the HSI technique, clearer understanding of the diagnostic potential of the HSI segmentation technique itself could be retrieved by combining the HSI and Raman imaging into one system to remove the need for coregistration between these two imaging modalities, and/or by using the same tissue section for all data collection. These options and their benefits will be discussed more fully in the future work section, along with methods to diminish error contributions 1 and 2 from the list above which could be directly applied to the current system.

Below is a table showing the sensitivity and specificity of the HSI with k-means cancer segmentation for each cancerous sample, and the technique as a whole across all samples, with two diagnostic thresholds; 12% and 8% tile coverage classified as a cancerous tile, versus the consensus tile cancer coverage of the three histopathologists. The standard deviation for the sensitivity and specificity of the technique across all samples in given in the final row.

The two diagnostic thresholds selected are 12% and 8% area coverage of a tile by the cancerous segment. The former was chosen to maximise both sensitivity and specificity, as would be required if HSI were to be used as a standalone diagnostic technique. The latter was chosen to maximise sensitivity without losing so much specificity that the false positive level becomes impractical. The 8% threshold was also the closest (of the thresholds selected for a smooth ROC curve) to the size of one squamous cell, of which only one is required the wrong side of the basement mebrane to garner a diagnosis of invasive cancer. In a combination HSI and Raman system, the initial HSI-based segmentation must have high sensitivity to ensure all potentially cancerous regions are highlighted for further interrogation with Raman spectroscopy. As the results are discussed in more detail and compared to work carried out by other groups, it seems prudent to highlight that the sensitivity and specificity of a diagnostic technique can

broadly refer to two different things. Firstly, to its potential to classify whole samples as containing, or not containing cancer. Secondly, to the proportion of a single sample which is correctly identified as cancerous or non-cancerous tissue. The former is more common but, as we are not yet training a diagnostic classifier but instead determining the worth of pursuing a certain technique for that purpose, this thesis will adopt the latter, more stringent definition of sensitivity and specificity henceforth.

Sample no.	12% Sens (%)	12% Spec (%)	8% Sens (%)	8% Spec (%)
9806	80	67	84	62
2217	58	86	74	61
4644	84	15	97	9
2577	33	43	61	15
9890	89	39	91	33
4734	22	81	43	62
All samples	66±28	66±27	79±20	51±24

Table 6: Table showing the sensitivity and specificity of the HSI and k-means cancer segmentation for each cancerous sample, and the technique as a whole across all samples, with two thresholds; 12% and 8% tile coverage classified as a cancerous tile, versus the consensus tile cancer coverage of the three histopathologists. The standard deviation for the sensitivity and specificity of the technique across all samples in given in the final row.

As this is the first body of work which specifically probes the (tongue squamous cell carcinoma) diagnostic capacity of (the k-means clustering segmentation of PCA dimensionality-reduced) hypercubes utilising visible and NIR light, collected on unstained tissue sections, there are no fully relevant comparisons to be made. Many other visible and NIR elastic light scattering techniques used to diagnose oral cancer have achieved excellent diagnostic performance, though in every case the hyperspectral data has been collected either *in vivo*, or in *ex vivo* tissue which has not been sectioned. Both of these sample types contain blood which has proven high discriminatory worth in cancer detection and which produce sharp and intense absorption features with high classification power.[79]

The extensive work of the Fei group falls within this category, achieving AUROC scores as high as 0.88 in diagnostic tests across samples from multiple patients, as in this study. In tumour-normal interface tissue, they achieved a diagnostic sensitivity, and specificity of 66 and 76%, respectively, where their definition of tumour-normal interface tissue applies to all of the samples used in this study. Their comparison to the gold standard histopathological diagnosis was also performed in a tile-wise fashion rendering its methodology very similar to this work.[155] In the relative absence of large quantities of highly cancer-correlated, absorbant haemoglobin, the performance of our

HSI technique with k-means clustering segmentation has exceeded expectations.

Our lower specificity score is expected as the source of absorption spectral features can be specifically identified whereas diffuse scattering signals are comparitively non-specific. The formidable diagnostic performance achieved by hyperspectral imaging by Fei *et al* was achieved using a Maestro (PerkinElmer Inc., Waltham, Massachusetts) imaging system which is significantly more expensive than the cost of our system, though its speed of data collection will be much higher due, in part, to the use of a non-pulsed Xenon source. The skeleton methodology of the work of Fei *et al* is very similar to this study, though the use of fresh excised tissue sections yields very different spectra and discriminatory spectral features, and would slot less efficiently into the current gold standard diagnosis workflow. Nonetheless, their work on incorporating spatial hypercube features inspired our investigation into utilising texture analysis as a potential segmentation and diagnostic aid in our own system, on unstained frozen tissue sections.

No studies utilising vis-NIR diffuse scattering spectra to discriminate oral cancer from thin, unstained tissue sections were found in the current body of literature. However, infra red microspectroscopy has been used successfully in this way. Long acquisition times plagued the early FTIR-based diagnostic systems, with a full tissue area scan taking several hours: almost as long as Raman spectroscopy. Kuepper *et al* exploited the growing availability of quantum cascade lasers to create an infra red microspectropy capable of detecting Stage II and III colorectal cancer with 96% sensitivity and 100% specificity compared to histopathology, in 30 minutes or less and with less than an hour of analysis time per tissue section.[156] This performance is comparable to Raman spectroscopy and significantly faster, and has potential to be an ideal solution to the growing pressure on UK histopathologists. It is, however, an expensive system containing both a costly microbolometer focal plane array (FPA) detector and a quantum cascade laser, known for their clinically-prohibitive expense.[157] IR detectors tend to limit the spatial resolution of the technique, though improved resolution can be achieved at a significant cost. The work by Kuepper *et al* used a microbolometer with $4.25 \times 4.25\,\mu m$ spatial resolution, half as fine as our 2 x 2 µm pixel size. This comparison is only significant in comparison to an HSI standalone system, rather than the HSI-Raman combination system we are primarily recommending which would limit the system's spatial resolution rendering a high resolution pre-screening process obsolete. Unfortunately, until further developments in semiconductor science, build costs will likely prohibit its clinical translation and widespread use.

Sample-Wise Performance Evaluation

The performance of the HSI, PCA-based dimensionality reduction and k-means

clustering technique varied significantly between individual samples in ways that can be tied to the specific limitations of that dataset (Table 5). Identifying connections between the performance of the HSI technique and specific sample or dataset challenges allows us to see the potential of the system once these sample-based limitations are minimised. Some potential solutions for these limitations are also suggested, while considering the total cost, speed and practical implementation of this approach.

The greatest limitation of the tilewise scoring evaluation technique is differences (both artefactual and real) between the fixed, H&E-stained tissue sections and the frozen, unstained HSI/Raman tissue section. Tongue squamous cell carcinoma sample **#9806** (shown in Figure 55) includes examples of all main types of sample errors/differences:

1. Severing (Either complete, or with rotation) (top right)
2. Folds (bottom left)
3. Change in porosity/holes between sections (in main cancer area, marked green)

While the ideal solution to this would be utilising the same tissue section for all imaging modalities by performing HSI and Raman on a frozen section which is stained with H&E later for histopathological evaluation, this comes with several practical difficulties. Instead, simple software solutions were employed to ignore areas with large differences between H&E and HSI/Raman tissue section such as severed/folded/rotated areas. However, very porous or damaged tissue is difficult to correct for as can vary dramatically between sections and greatly affect individual tilescores across an otherwise uniformly "cancerous" region. Tiles in these areas could be automatically excluded, but in many cases this would leave insufficient tiles for robust performance evaluation.

The aforementioned separated and rotated area of sample #9806 contained several lobes of inflammation and, had it not been automatically excluded as coregistration was not possible, would have decreased the specificity score of the HSI technique for this sample.

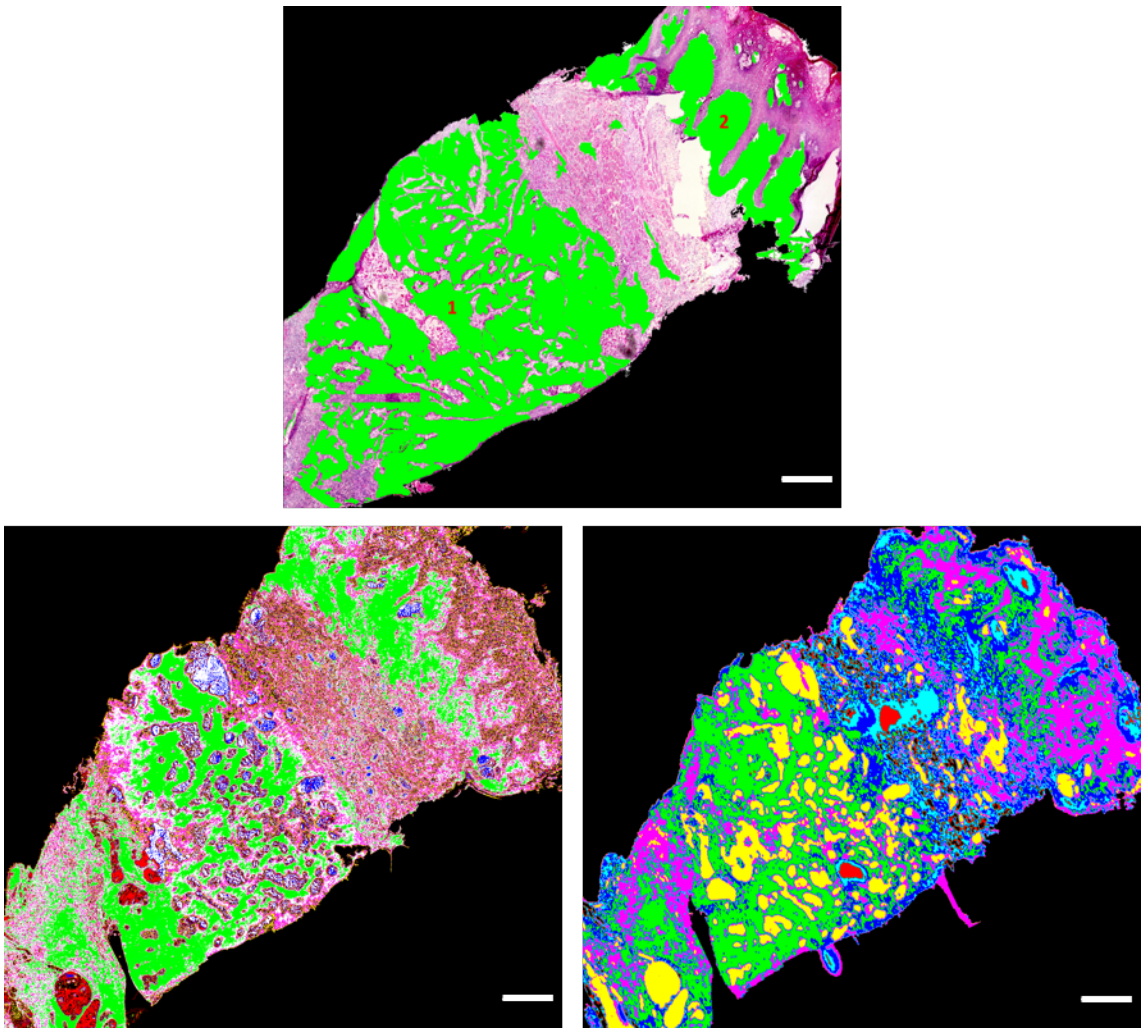

Figure 55: Comparison of cancer segmentation afforded by histopathologists (1 = cancer consensus between three histopathologists, 2 = inflammation demarcated by one histopathologist), dimensionality-reduced and k-means clustered HSI and Raman spectroscopy (top to bottom, left to right) of tongue SCC sample #9806. Scale bar represents 500 µm.

When imaged by HSI, the areas of inflammation are separated into the same cluster as cancer itself which suggests that HSI cannot reliably separate cancer and inflammation. The inflammation was only selected by one histopathologist on three of the six samples so there is no quantitative accounting for interobserver variation, nor sufficient sample repeats for reliable assertions.

There are, however, similarities between some inflammatory cells and cancer cells which could affect the scattering spectra similarly given that ESS has been been shown to be sensitive to subcellular changes in optical index of refraction.[158] Common features of inflammatory cells, and cancer cells, primarily occur in relation to the nucleus, which is known to be the source of several important discriminatory features in elastic-scattering based cancer diagnostics including nuclear size, pleomorphism, increased

levels of chromatin and increased nuclear-to-cytoplasmic ratio.[159] For example, hallmark multinucleated giant cells of chronic inflammation could exhibit similar elastic scattering effects to a cancer cell which commonly contain large, pleomorphic, unusually shaped nuclei.[147] This confusion between inflammation and cancer has also been reported by other HSI research groups, including Jerjes *et al* who hypothesised this could be to blame for the high false positive rate in their visible range ESS-probed fixed oral squamous cell carcinoma samples.

The body's natural inflammatory-based defence response at the tumour margin further compounds the discrimination challenge at the most important interface; cancer lesion and the sound host tissue. An over-estimation of cancer would be safer as more would be resected giving clearer margins. However in critical sites where no spare margin exists, such as when cancer impinges on a vital structure like nerves or blood vessels, this could risk a decision for inoperability and palliation where a cure was possible. This disadvantage of the HSI technique could be significantly mitigated by including Raman spectroscopy in a combination technique. Unlike ESS based techniques, Raman spectroscopy is chemically specific and is therefore less susceptible to confusion between these two cell types and has been proven to be able to discriminate between inflammation and cancer.[160] This aligns with the comparison figure above, where the Raman map area corresponding with the inflammation selected by the histopathologists is separated into multiple segments (green cancer, royal blue and blue).

Attempts to quantify this confusion identified that the three samples on which inflammation was demarcated by a histopathologist happened to have higher sensitivity and specificity than other samples. Therefore, the following ROC curves in Figure 56 were evaluated only with those three samples, #9806, #2217 and #4644 to avoid bias.

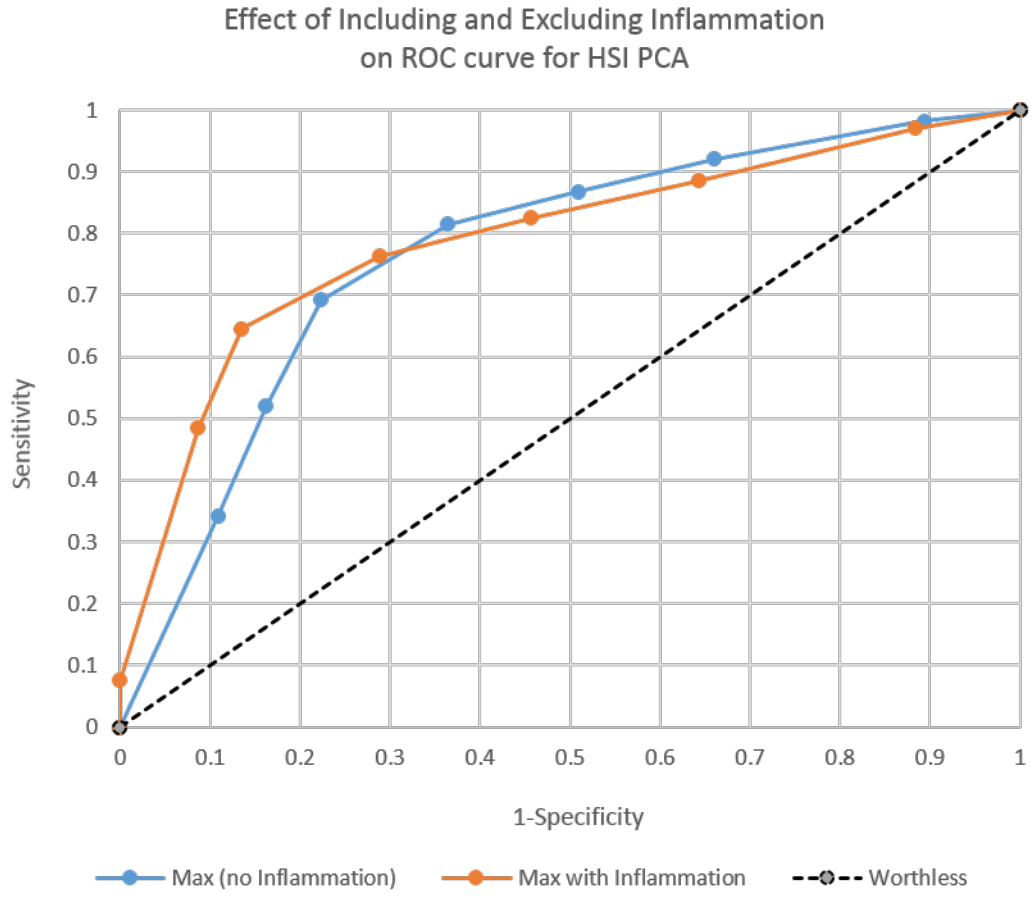

Figure 56: ROC curve for max of all histos with and without inflammation. (just samples 9806, 2217 and 4644)

As predicted, the HSI-based segmentation performed slightly better with respect to the gold standard when inflammation was included in a true positive tile diagnosis, with an AUROC score increasing from 0.77 to 0.79 in this case. Chronic inflammation has been shown to be an important factor in promoting tumorigenesis in oral cancer and therefore, its inclusion in the initial segmentation provided by the HSI in a combined HSI-Raman system may still be relevant in disease progression monitoring.[161] This line of enquiry should be expanded to include other samples and inflammation demarcation by additional histopathologists, but appears worthy of further investigation.

In contrast to the high sensitivity achieved by the HSI cancer segmentation of sample #9806, the sample which reported the lowest sensitivity was sample number #**4734**. This was one of the two samples which required four field of view tiles in order to capture the full tissue area. This means there is a large cross-shaped area of overlapping tiles, roughly centred to the tissue section. The area of overlap suffers from two potential sources of error: Firstly in the method chosen which selects the maximum at each pixel in overlapping area which will, by nature, accept a higher proportion of bright pixels and

outlier pixel values such as those from dead pixels. This increases the average grayscale intensity and decreases the SNR in these regions, relative to the rest of the tissue area. This decreases the likelihood of the same segments being consistently selected across overlapping and non-overlapping tissue areas. Secondly and similarly, if overlapping fields of view are not of similar brightnesses in these areas, the k-means segmentation cannot identify common segments across their joining boundary, breaking up a segment of one feature into multiple, and under-representing the sensitivity of the HSI technique to cancer in that sample.

In order to mitigate the latter source of tiling error, the white light laser source utilised in Setup 2 (7.1) was replaced with an incoherent pulsed Xenon light source. This facilitated a complete flat-field background correction protocol by removing the shifting squircle background which prevented this and plagued the data in Results 2 (10).

In addition to this background correction, an ImageJ macro was designed to perform fine corrections to match the brightness of subsequent wavelength slices to each other (9.2.1). In most samples this generated a significant improvement in tiling artefacts, even with the continued use of the maximum pixel value tiling method. In some samples, there was very little suitable background to select an area for the ImageJ brightness matching macro which rendered its performance worse.

Additionally, there were two samples, #4734 and **#9890** which required twice as much tiling as the others and which therefore suffered in quality and reported sensitivity. The large overlapping regions of #4734 were particularly problematic due to the relative position of the cancerous region, whereas in #9890, the off-centre cancer segment was less affected. The poor SNR of the overlap areas in **#4734** meant that segments were disjointed and sparse, as was the case with the selected cancer hue. Due to the absence of sizeable areas of connected segments, the sensitivity of the HSI imaging process on this sample was low and, in a sense, under-reported. This did, however, translate to relatively high specificity as the small and sparse cancer segment was almost completely within the bounds of the gold standard histopathologist-selected cancer area. The multiple folds in the H&E stained tissue section, which can be seen in Figure 57, also affected the quality of the gold standard diagnosis as affine transformations cannot reproduce a folding effect on the HSI data to perfectly coregister the two datasets. Even if the HSI and histopathology selected cancer segments matched perfectly, this folding would decrease true positives and true negatives by requiring compression of the HSI dataset to "match" the folds, decreasing the reported sensitivity and specificity. The effect on the specificity was diminished again in this case by the very small cancer segment selected by the HSI inside the cancer region selected by the histopathologists.

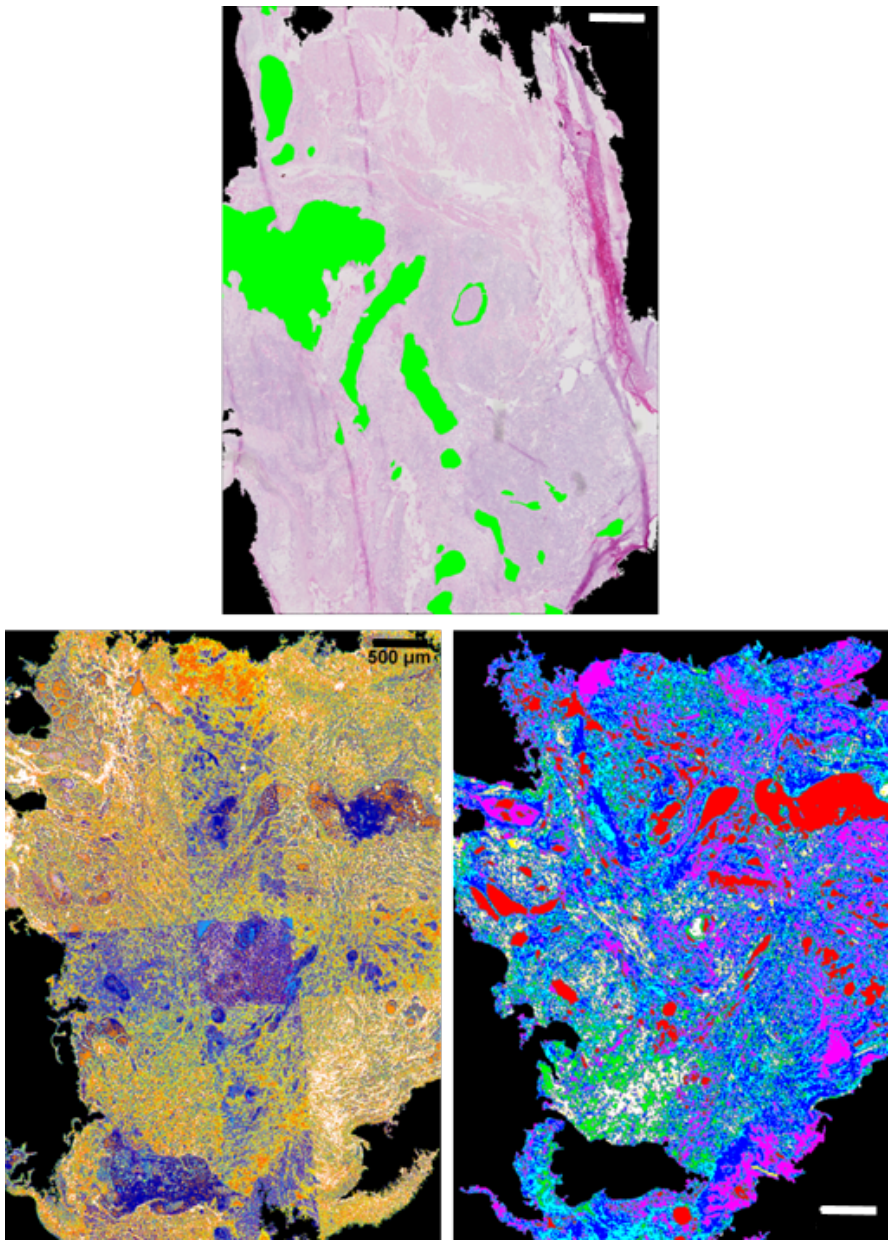

Figure 57: Comparison of cancer segmentation afforded by histopathologists (consensus between 3, shown in green), dimensionality-reduced and k-means clustered HSI (royal blue) and Raman spectroscopy (white) (top to bottom, left to right).The k-means clustered map of tongue squamous cell carcinoma sample #4734.

When we add the Raman data into the comparison, our evaluation of this sample changes again. There are no segments in the Raman which appear to match the gold standard diagnosis well. This means either that the Raman cancer segmentation performed poorly on this sample, or that the position of the cancerous tissue regions changed dramatically between tissue sections. The average of five Raman spectra sampled from the the white segmented cancer region in the Raman area map were plotted, along with five spectra sampled from elsewhere in the sample (diagnosed as

non-cancerous by the histopathologist consensus). The ratio of the 1455 (related to relative proportions of lipids and proteins) and 1655 cm^{-1} (C=O stretching of collagen and elastin) Raman spectral peaks have been used as a quick diagnostic test for several varieties of cancer, with a value of > 1 being generally accepted as the healthy tissue range and <1 indicating cancer.[162] In the white segmented area of the Raman scan, this ratio comes to ~ 0.98, just meeting the criterion for cancer. This implies that the Raman cancer segmentation may be correct and a comparison between the Raman and HSI datasets is more diagnostically relevant in this instance. A higher sensitivity score of the HSI cancer segmentation compared to the Raman segmentation might therefore be expected, however this is not the case with 18% and 38% sensitivity reported for the 8 and 12% diagnostic thresholds, respectively. Visual comparison, bearing in mind the potential for hue-switching across tile boundaries, shows some similarities between the Raman and HSI segmentation, particularly in the central island of cancer (royal blue in the HSI segmentation map) and bottom left (orange segment in HSI). Unfortunately the current inter-modal comparison process cannot take this hue switching tiling error into account like the human brain can.

Interestingly, the tissue holes selected in red in the Raman map are frequently included in the cancer segment in the HSI. Where forward Mie scattering contributions are hypothesised to increase in areas of high nuclear density and size, an increased intensity across all wavelengths would be expected. In HSI these regions could report similarly to image holes which would also exhibit higher pixel intensities due to relatively uninhibited transmission of light through the calcium fluoride slides. This confusion could also explain why we do not see the improved sensitivity we expect when we compare the HSI to the Raman, in spite of their observable similarities. Additionally, the Raman datasets are more prone to coregistration errors (as discussed in 15.2.4) which could decrease reported sensitivity, especially in the case of very small selected cancer areas such as in this sample. Re-collecting the HSI data for this sample at a lower magnification with little or no tiling would be an interesting exercise and could help illuminate the contribution of the tiling errors to the performance evaluation. This is discussed further in the Future Work section (16).

The sample which reported the lowest specificity in its cancer hue selection was **#4644**. This sample contained one central region of good quality in both the H&E stained section and the frozen section with one small area which is detached and translated, and one large hole with neighbouring tissue folds. Unfortunately, a very large proportion of this high quality sample region contained cancer and the healthy tissue was mostly present at the tissue edges, leading to its exclusion by the tile selection macro (14.2.5). This dramatically reduced the sampling of true negatives and decreased

the reliability of the specificity evaluation.

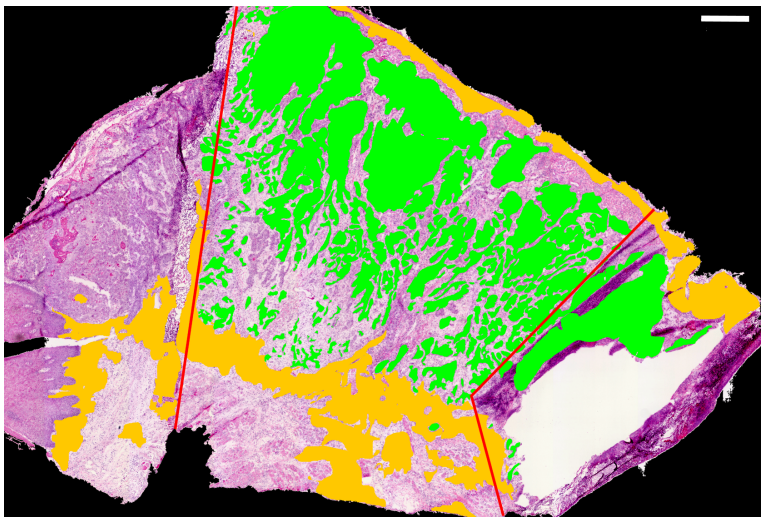

Figure 58: The RGB image of the H&E stained squamous cell carcinoma tongue tissue sample, #4644, with concensus histopathologist cancer demarcation in green, the inflammation demarcated by one histopathologist (orange) and areas of damaged or translated tissue areas separated by red lines, with only the central strip being considered for comparison.

Earlier in this section, it was hypothesised that the HSI technique was not sensitive to the differences between cancer and inflammation and therefore struggled to discriminate them. If this is the case, the area of tissue which would garner a "healthy" tissue diagnosis is smaller still (Figure 58), further decreasing the reported specificity.

The lowest overall performance of the HSI segmentation technique was on #**2577** which suffered from the largest selection of specific limitations of any sample (Table 5). An already small sample at < 7 mm in its largest dimension, large portions of the datasets were rendered unusable due to the high proportion of fragmentation that occurred in the frozen section which can be seen clearly in Figure 59. This small area of inclusion decreased the reliability of sensitivity and specificity measures for this sample. Coregistration was also complicated, with a higher probability of human error in the selection of the common coordinates, for use in the affine transformation, across the three different datasets.

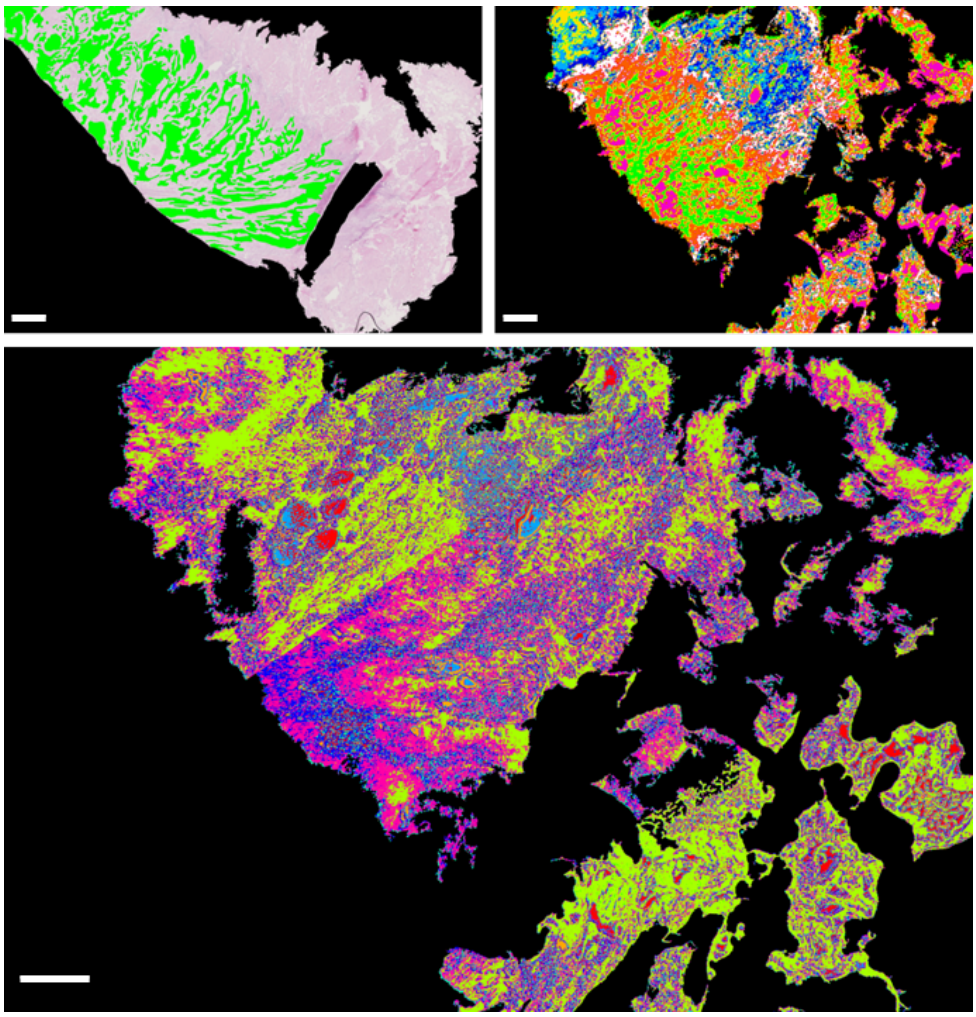

Figure 59: Comparison of cancer segmentation afforded by histopathologists (consensus between 3, shown in green), Raman spectroscopy (pink) and dimensionality-reduced and k-means clustered HSI (pale blue) of tongue squamous cell carcinoma sample #4644 (left to right, top to bottom). The scale bars represent 500 µm.

Again, the tiling artefact has selected what appears to be the incorrect cancer segment due to the hue-switching across the overlapping field of view tile boundaries. While the new background and brightness correction protocols have decreased the number of samples which are affected by tiling artefacts in comparison to Results 2, it was still a substantial source of error in three of the six samples. Writing the HSI data collection software on a platform which facilitates the triggering of the pulsed Xenon source by the camera would help to stabilise the brightness at each wavelength across multiple datasets. A platform with this capability such as LabView may also allow disabling of the rolling shutter in preference of a global shutter. Without the correction step the rolling shutter required, which involved taking multiple datasets and then taking the maximum pixel value at each position across all repeats, less noise would be introduced into the HSI data as a whole. This would be a prudent consideration for future work

on this system.

The selected Raman cancer hue (pink) appears to correlate with holes in the sample, suggesting it is incorrect. The orange and lime green hues appear to overlap more with the cancerous region selected by the histopathologists. Inspection of the Raman spectra in the orange and green regions suggest that the former is the cancerous tissue: Along with protein-correlated peaks in both segments, Raman spectra from the orange segment reported the presence of guanine, adenine ring breathing.[54] Where the protein increase might suggest hyperplasia or dysplasia, the presence of additional spectral contributions from DNA implies aneuploidy and/or multinucleation which is characteristic of cancer cells. The low Weber contrast of the Raman data prevented the identification of more clear spectral peaks.

This low Weber contrast also implies a low SNR, which differs from Weber contrast in that it is a measure of the "contrast" of only the desired signal component relative to noise or background. Without having identified specific diagnostic classifiers, identifying the desired signal component is not possible and the SNR cannot be calculated explicitly. When trialling various light sources and monochromator slit widths, observing the transition between meaningless and meaningful tissue segmentation facilitated an assertion that the SNR of both modalities is low. Nonetheless, in order to further characterise the system performance and to chart changes to its performance with system modifications the SNR of both the HSI data and the Raman data will be calculated once diagnostic classifiers have been selected.

Cancers which have been sectioned at an angle to their infiltration often exhibit these finger like islands of tumour, rather than a large body of cancer such as in sample #9806, pictured in Figure 55. These islands are likely to look very different in subsequent tissue sections, in spite of occupying a similar region of the tissue. In cases like this, if the outer boundaries of the cancer island-filled region were selected as the gold standard diagnosis we would more accurately select corresponding segments in other measures by the overlap coefficient method (14.2.4). We would also retrieve more representative performance measures. Though the matter of exactly how to define the edges of such regions for delineation is non-trivial.

The pale blue cancer segment selected as cancer in the HSI dataset, potentially erroneously, is difficult to see in the figure as it is diffusely present over most of the top left region of the tissue, to the left of the erroneous tile hue switch. The sparseness of this segment would contribute to its low sensitivity score, similar to sample #4734. If we instead assume that the orange and green segment in the Raman area scan is malignant as a result of the evidence presented above, then we begin to see some similarities between the Raman and HSI data. For example, to the right of the tile hue switch in the HSI data, the royal blue and pink segments match the green and orange

Raman segments well, implying hyperplasia or dysplasia, and cancer respectively.

The HSI tissue segmentation performed well on sample #**2217**, but there were some sample artefacts which became apparent in the later stages of analysis and which can be seen in Figure 60. Two very fine cracks in the very fragile calcium fluoride slides caused a slight separation between the slides where part of the tissue section was adhered to the top slide, and some to the bottom. In the Raman spectroscopy area scan of this sample, due to its confocality, only information about the top layer of the tissue is reported. In the data collected with the transmission-mode HSI system with the high field depth objective lens, some information is retrieved from all parts of the tissue, though features from the different layers of tissue have been segmented separately even though they may be of the same pathology.

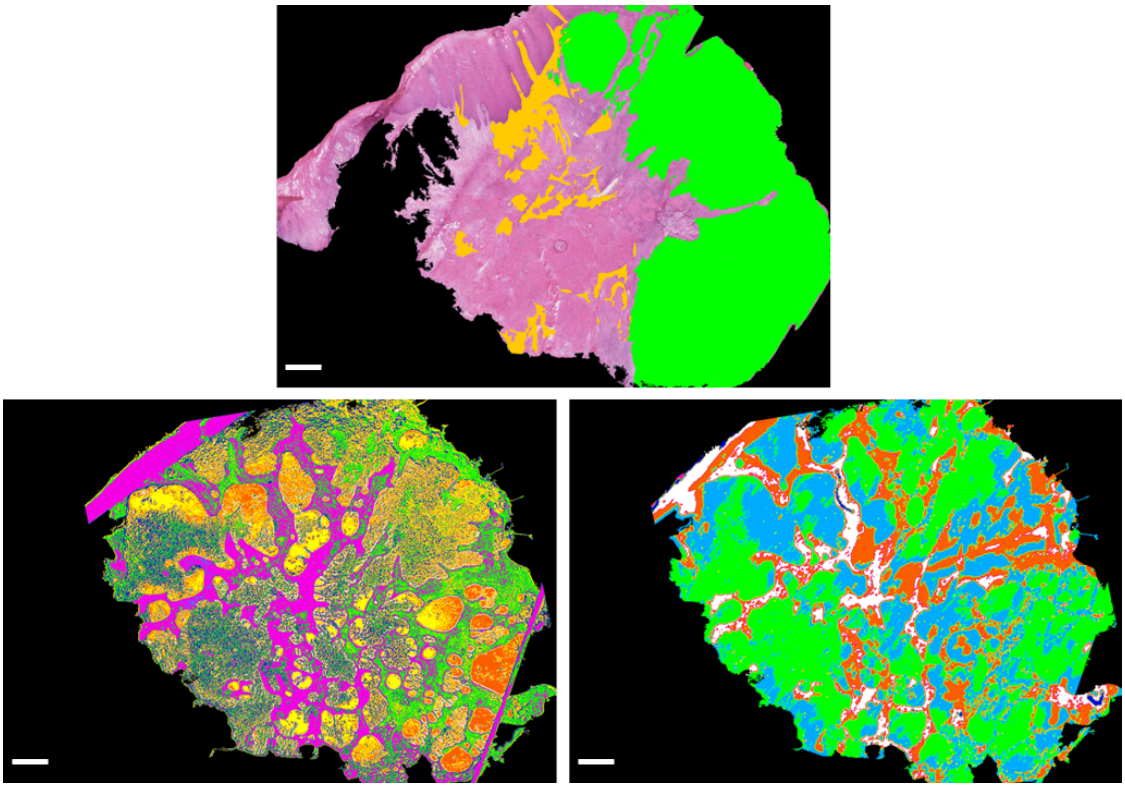

Figure 60: Evidence of the slide crack-induced artefact in tongue squamous cell carcinoma sample #2217. Cancer segmentation afforded by histopathologists (consensus between 3, shown in green) and inflammation segment by one histopathologist (orange), dimensionality-reduced and k-means clustered HSI and Raman spectroscopy area scan (top to bottom, left to right). Scale bars represent 500 µm.

The orange and white regions in the Raman equivalent (bottom right of Figure 60) appear to show the tissue adhered to the top cover slip and can be correlated with the magenta and green regions in the HSI k-means cluster map (bottom left of Figure 60). They differ slightly in the areas which they occupy, especially in the bottom right of

the tissue area, with the Raman detecting a smaller area as expected by its filtering out of out-of-focus contributions which is not a feature of the HSI system. The other hues in the Raman map, green and blue, are randomly distributed across the entire background and therefore do not reflect useful information gleaned from the tissue itself. On the other hand, the sparse yellow and mustard yellow segments in the HSI map can be seen to concentrate in the cancerous and inflamed regions indicated by the gold standard diagnosis, commonly co-occurring with the green segment highlighted in the foreground tissue. This co-occurence implies that both segments, and therefore both the foreground and background tissue artefactual "layers", contain important diagnostic information. As these two tissue layers are forced into separate clusters by their artefactual separation and the associated changes to their reported spectra, the sensitivity of the HSI technique's performance on this sample is likely under-reported relative to the gold standard histopathology-based diagnosis.

Due to to its larger area, the mustard yellow exhibited the highest overlap coefficient with the histopathologist-selected cancer region and was therefore selected as the cancer hue rather than the green. With information from this deeper tissue area not accessible by Raman spectroscopy, the cancer segments from both the HSI and Raman data agree poorly with each other. This is can be seen clearly if we calculate the sensitivity and specificity of the HSI performance on this sample, taking the Raman cancer region as the gold standard: It performs with sensitivity and specificity of 34 and 63%, compared with 58 and 86% with respect to histopathology. Using an objective lens with an even lower numerical aperture would further increase the depth of field, ensuring samples with a range of slide separations could still be segmented by HSI, though this would only be useful in a standalone HSI system, not a HSI-Raman hybrid system. As a lower magnification objective is already a suggested solution for minimising tiling artefacts this could be explored in tandem.

Adaptations for Clinical Application

The use of an objective lens with a lower magnification and NA could be accompanied by further potential benefits to the industrial application of this system: increased speed of data collection due to shorter required exposure times provided by increased light acceptance and increased analysis speeds due to much smaller datasets. Additional ways to reduce run time include reducing wavelength range or step size. The PCA data collected from our datasets to inform their dimensionality reduction also assist us in identifying which wavelengths contribute most important information and which wavelengths could be cut out of the hyperstack collection.

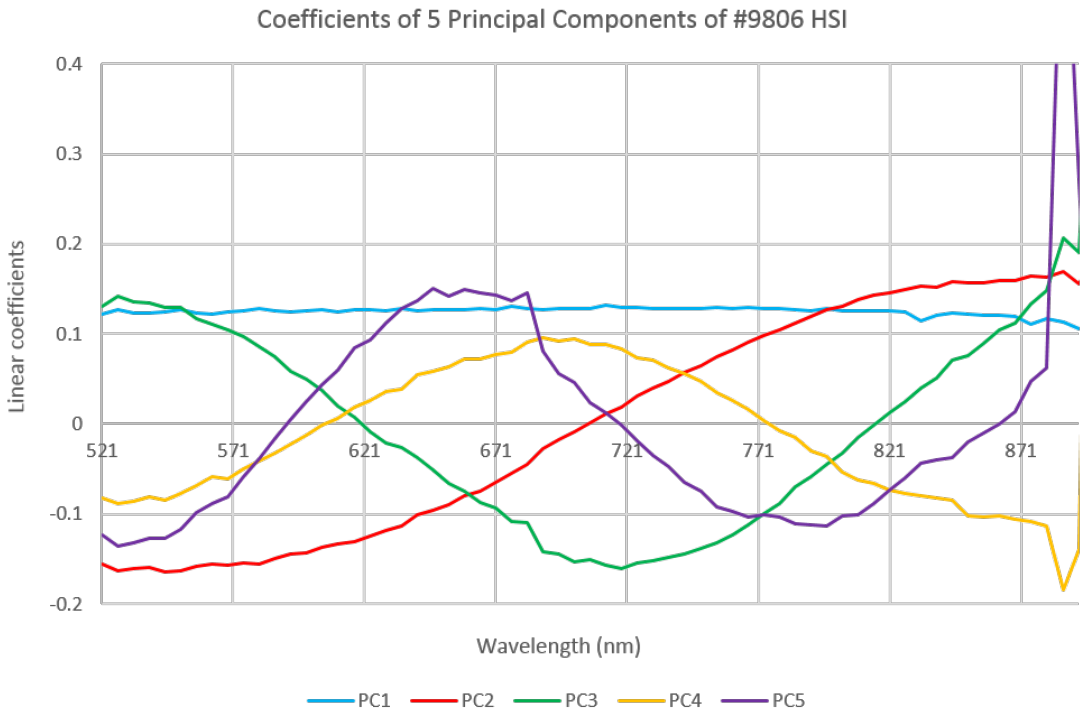

Figure 61: Coefficients against variable (wavelength) plot for the tongue squamous cell carcinoma sample #9806 hypercube.

Each principal component is a linear combination of of all variables which contribute to the data: In our data, the variables correspond to the different illumination wavelengths. The larger the magnitude of the coefficient, the more important the corresponding wavelength is in calculating that component, and its sign indicates positive or negative correlation of that compoent with that wavelength feature. This relative component importance and the correlation direction is depicted with a graph of coefficient against variable in Figure 61.

PC1, which describes over 80% of the variance of the datasets from every sample, is described by a remarkably evenly weighted linear combination of every wavelength in the range. This is in keeping with many other studies which failed to identify more valuable wavelength regions.[163] Mie scattering across this wavelength range for scatterer sizes of ~ 3 µm and ~ 9 µm corresponding with healthy cell nuclei and cancerous nuclei (chosen with reference to HeLa cell data and [69]) differs most at our shortest wavelengths, with decreasing difference as the wavelength increases, as shown in Figure 62.

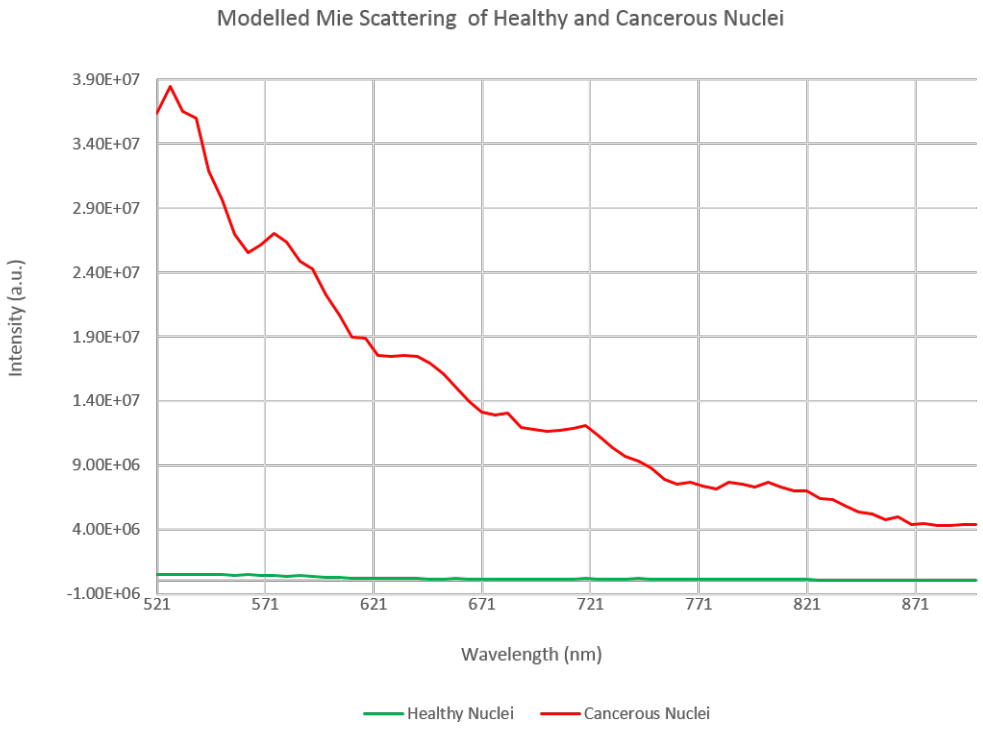

Figure 62: Plot of Mie scattering (at 0^{o}) intensity against wavelength for monodispersed healthy size nuclei (~ 3 µm) and cancerous nuclei (~ 9 µm) using Philip Laven's open source MiePlot software. (Air was selected as the material either side of the sample, though calcium fluoride was tested and yielded a similar enough trend for the purposes of this discussion.)

This decreasing scattering intensity with increasing wavelength means that the best Mie scattering-based discrimination occurs in the same spectral regions which contain potential contributions from residual haemoglobin, discussed below. This wavelength dependent component of Mie scattering may, therefore, be obscured by more intense and distinctive haemoglobin absorption features. Discriminatory value gleaned can still be gleaned from Mie scattering as the intensity of scattering is consistently higher from large, cancerous nuclei than smaller, healthy nuclei: This could be the primary source of principal component (PC1).

While the order of importance of subsequent principal components varies between different samples, PC1 is always comprised of an approximately even contribution of information from each wavelength. (For this reason, the principal component plots were not average across all samples and the plot above is just for samples #9806, though these important wavelength regions are present across all samples)

PC2 has increasing contributions above 700 nm until the upper limit of the wavelength range (899 nm) which correlate with the scattering properties of fibrous tissue components such as collagen.[164] Collagen and elastin fibres change in many ways as

tissue progresses from healthy to neoplastic, including decreasing in order and packing density.[165] We would, therefore, expect negative coefficients at these wavelengths as the increasing sparseness of fibres in oral cancerous tissue should lead to diminishing spectral contributions in this range. This is observed in PC4 which is a less influential component than PC2 which shows the opposite: positive correlation of data variance explained by spectral features in the fibre-scattering-encoded wavelength. This increasing sparseness and diminished spectral contribution of fibres in oral cancerous tissue, however, could be attributed to an increase in Type III collagen, which primarily occurs in single fibres rather than bundles.[166] This potential increase in Type III collagen would likely occur in tandem with a enzymatic degradation of Type I collagen which could explain this dichotomous result.

The apparent influence of data collected at high wavelengths could also be an artefact of decreasing SNR due to the lower efficiency of the optical components and camera sensor as wavelengths increase above ~ 850 nm.

PC3 shows a peak around 540 nm, with high coefficient numbers from the starting wavelength (515 nm) to approximately 580 nm. This range encompasses every potential spectra feature which could be produced by residual haemoglobin in the samples; absorption at 540 and 580 nm for oxygenated haemoglobin and at 560 nm for deoxyhaemoglobin. A change in the balance between oxygenated and deoxygenated haemoglobin could lead to a broad contribution of discriminatory power between these wavelengths, as exhibited in the principal component plot. This discriminatory power could also be bolstered by the probable increase in blood residue present in cancerous tissue samples due to angiogenesis in cancerous tissue, instigated by chemical signals produced by the tumours themselves to provide additional oxygen and nutrients to accommodate upregulated cell proliferation and waste production.[167] These wavelengths may also contain important scattering-based spectral features, as ~550 nm was also identified as the most important discriminatory wavelength by Khouj *et al* in their work on segmenting unstained, formalin-fixed ductal carcinoma tissue sections.

Given that the samples are colourless themselves and unstained, we may expect Mie scattering effects to be a primary contributor of diagnostic information in all cancerous samples with the proportion of fibrous tissue and residual haemoglobin varying more dramatically between samples according to inter-patient variation, biopsy location (even within the tongue), and small differences in sample preparation. Multiple principal components in this plot show combinations of features at the highest and lowest wavelengths which we have tentatively assigned to blood and fibrous spectral contributions. In cancerous tissue, we expect some increases (deoxyhaemoglobin and Type III collagen, for example) and some decreases (Type I collagen and oxyhaemoglobin) of tissue components with spectral features in both of these wavelength ranges. There-

fore, multiple principal components with different combinations of positive and negative coefficients in these regions are to be expected.

The three main peaks displayed by PCs 2 - 4 were used to select three wavelength regions, 521 - 635 nm, 605-755 nm and 815 - 899 nm, which are weighted more heavily in their contribution to these principal components. It is worth noting that, between them, they span almost the entire wavelength range which supports the assertion that all wavelengths are important for discriminating oral cancer from healthy tissue using elastic scattering-based hyperspectral techniques. Nonetheless, with the aim of improving the speed of the system, three sub-stacks of the aforementioned wavelength ranges were generated from the 9806 HSI data and segmented with k-means clustering for qualitative evaluation (Figure 63).

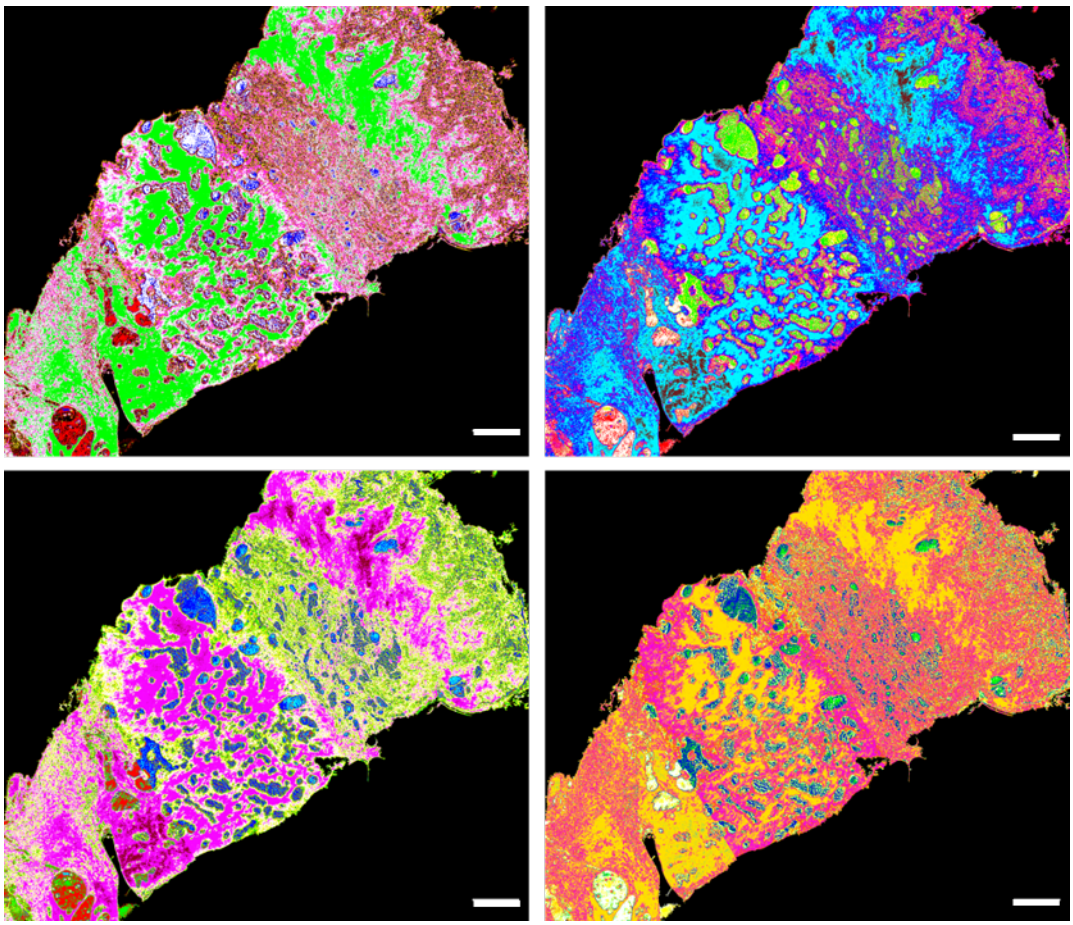

Figure 63: False coloured k-means clustered HSI of cancerous sample #9806. From left to right, top to bottom with cancer hue indicated in brackets: Full wavelength range 521 - 899 nm, 5 principal components (green), 521 - 635 nm (light blue), 605-755 nm (magenta) and 815 - 899 nm (orange), all principal components. Scale bar represents 500 µm.

The presence of both main segments (orange and magenta) across the entire tissue area in the highest wavelength hypercube (815 - 899 nm) evidences less specific seg-

mentation. Additionally, tiling errors are more prevalent as can be seen at the bottom left of the main cancer region where the hue switches from magenta to orange across the tile overlap area boundary. Both of these outcomes could be due to the lower SNR in these wavelengths due to system instability and lower illumination power.

The resulting increase in erroneous pixels was also exacerbated by their preferential acceptance both in the rolling shutter correction macro and in the maximum-pixel-value tiling method. Due to the random nature of this noise, false discriminatory spectral features will be present in this wavelength range, and inseparable from real diagnostic features such as a general signal increase due to decreased attenuation by the tissue. In the current permutation of the system (Setup 3) and its associated software, this wavelength range alone does not provide the level of segmentation specificity required from a pre-screening process to be used in tandem with Raman spectroscopy for oral cancer diagnosis, though it does add some value in a full-wavelength hypercube.

On the other hand, the first two wavelength ranges appear to perform very similarly to the full wavelength range, though with some minor tiling based errors visible in the bottom left of the the main cancer area. There appears to be a slight increase in random noise which could be due to the absence of dimensionality reduction by PCA in this section of the investigation. Interestingly, they both appear to provide additional segmentation within the lobes of inflammation at the top right of the tissue, compared to the full hypercube. The shapes and patterns of the segments here can be visually correlated with the corresponding region in the Raman dataset in Figure 55, suggesting this may be related to the separation of inflammation from cancer.

When the inflammation study is expanded for significance, this potentially wavelength-dependent inflammation segmentation would be an interesting line of enquiry to pursue in tandem. Quantitatively comparing the performance of the system over these two reduced wavelength ranges by calculating tile scores and comparing them with the consensus histopathology diagnosis, with and without dimensionality reduction by principal component selection, would be a worthwhile future endeavour in the development of this combined system for clinical application.

15.2.5 HSI Texture Analysis and K-Means Clustering

Preliminary testing of HSI on HeLA cells, with and without nuclear disruption with paclitaxel (15.1), suggested that spatial information may be important in separating regions with different scattering responses.

Work on automated feature extraction and ranking by Fei *et al* found that texture features extracted from the cytoplasm and nuclei of H&E stained oral epithelium were often ranked highly for discrimination of oral neoplasia from healthy tissue and cor-

relation with spectral HSI information.[102] Specifically they found that, in neoplasia, fractal dimension measures increased in tandem with reflectance, and the mean perimeter of Delauney triangles decreased. They hypothesised that these significant textural alterations were linked to loss of cellular organisation and nuclear consistency, along with nuclear crowding in malignant tissue. Nuclear crowding, in a transmission system such as ours, could lead to increased reflection away from the detector, and a resulting decrease in detected signal intensity. The texture metric "contrast" measures the magnitude of intensity differences between each pixel and those around it, making it suitable for reporting the regions of crowded nuclei associated with cancer. It may also report smaller crowding-like textural effects caused by increased nuclear inconsistency in these regions as ESS is sensitive to subcellular features, however work confirming this has ordinarily been performed at a significantly higher magnification than our system utilises, with many other groups utilising x40 magnification where we use x4 for speed of data collection for practical clinical application.

In the wavelength range reduction discussion, we identified that fibres such as collagen undergo changes in a cancerous environment in order to facilitate cancer cell infiltration. Therefore, texture alterations in collagen were also considered in the selection of specific texture metrics. Mostaco-Guidolin *et al* reported more significant detection of fibrous changes by GLCM texture metrics than first order statistics, bolstering our decision to look at local contrast. Their work showed that the homogeneity (or inverse difference moment) texture metric exhibited differentiating power between different collagen morphologies: Local homogeneity is higher in regions of organised and low contrast features, such as collagen in healthy tissue and healthy tissue itself, compared with the disorder which plagues both in neoplasia.

All of the standard GLCM texture metrics (contrast, correlation, homogeneity and energy) were calculated in tandem for each sample in Matlab but contrast and homogeneity were hypothesised to produce the most relevant differences between cancer and non-cancer (15.2.5). These texture image stacks were then segmented by k-means clustering, with no prior dimensionality-reduction with PCA.

Performance Evaluation

ROC curves plotted for each texture, and its segmentation peformance relative to the gold standard histopathological diagnosis, are not illuminating. AUROC scores are improved by the plotting of multiple different sensitivity and specificity pairs achieved by applying different diagnostic thresholds. Due to the patch-wise output of the texture analysis (designed to match the evaluation tiles), each tile is classified by the texture analysis as either 100% cancerous or non-cancerous. Unfortunately that means that the diagnostic threshold (the percentage tile area coverage by the cancer segment

which leads to a tile diagnosis of "cancerous") is effectively set at 100% and varying of the diagnostic thresholds for the calculations of true/false positives/negatives does not affect the performance outcome. The ROC curve is, therefore, mostly a straight line which adds little value in the performance evaluation of the texture-based segmentation, and substantially under-reports the AUROC score of the technique. A qualitative evaluation of the segmentation afforded by the texture analysis is sufficient to identify some of its strengths and limitations, and to suggest parameter alterations for improved future work.

The selected patch size (200 x 200 µm) meant that this analysis was sensitive to larger tissue features such as holes and cracks, which are unavoidable with such a delicate tissue preparation procedure. Often they are grouped into one segment with the healthy tissue, but occasionally they are too different in texture to any other tissue component or feature and therefore occupy an entire cluster of their own. As there are a fixed number of clusters, this "wastes" more meaningful segmentation opportunities. This would be difficult to overcome with a simple software solution due to the range of artefacts which might ellicit this clustering outcome. This is not the only common source of limitations which the texture analysis shares with the direct clustering of the HSI data. It appears that the tiling slightly depreciates the performance of the texture analysis too, but predominantly due to the increased random noise which affects the reporting of co-occurrence matrices and textures calculated from them. Fractal dimension, as utilised by Lu *et al* and Ou-Yang *et al* are less sensitive to noise, though they require either longer run times or more powerful computers, both of which a could not be widely afforded a benchtop clinical technique.

The contrast and energy maps showed more heterogeneous segmentation than the correlation and homogeneity maps, which rarely had tissue-wide assignation of more than 3 different clusters. This may mean that they are sensitive to tissue features other than, or in addition to cancer. The effective pixel size of this technique (200 x 200 µm) must be reduced if the more complex contrast and energy segmentation is to be attributed to real tissue features, with the help of histopathologists.

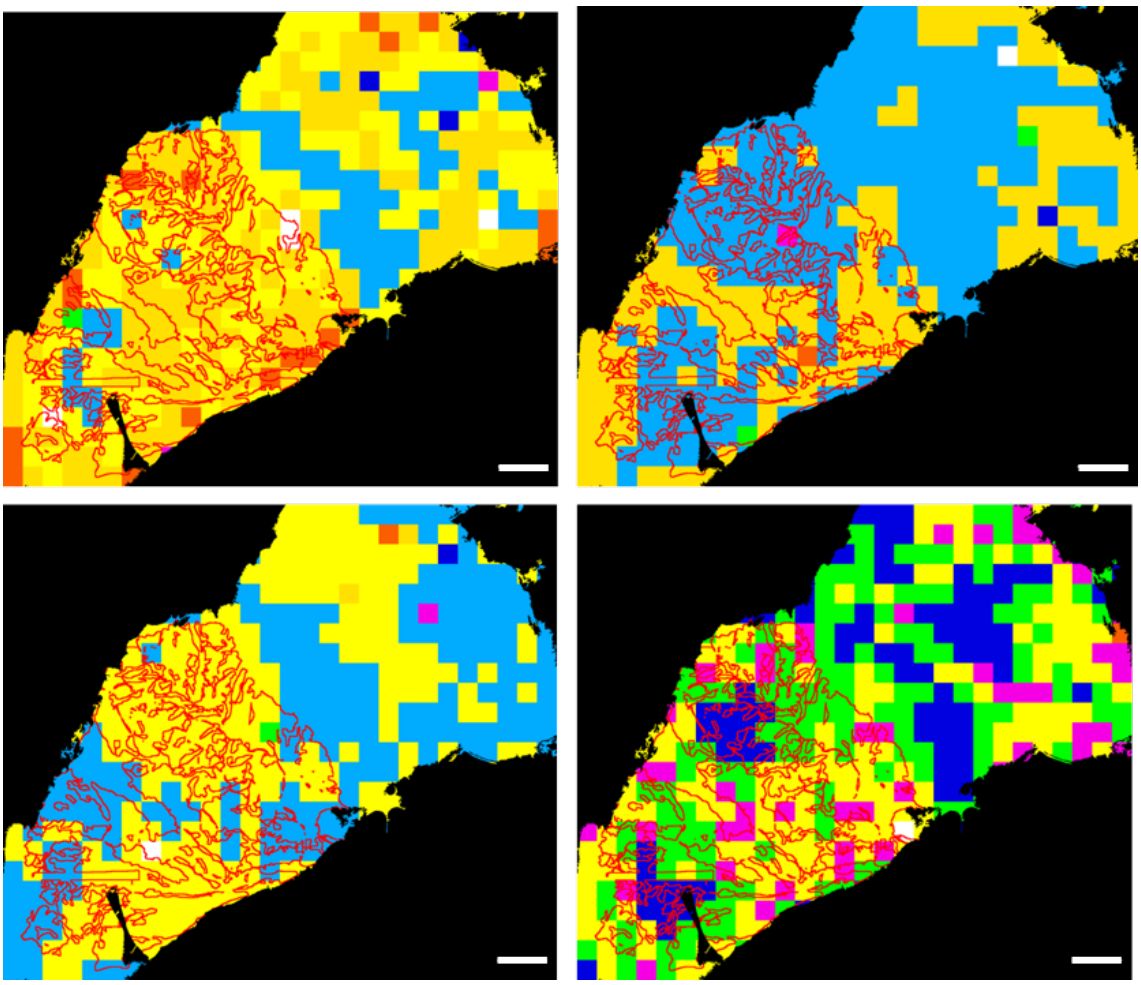

Figure 64: K-means clustered maps of the different texture analysis stacks produced from the HSI data from tongue squamous cell carcinoma sample #9806. Scale bar represents 500 µm. Top to bottom, left to right: contrast, correlation, homogeneity and energy texture metrics.

In sample #9806, the least relevant segmentation is produced by the correlation texture metric. Correlation is a measure of the dependence of a pixel's grayscale value on the grayscale values of the pixels around it. High correlation suggests one or multiple repeating patterns occur within the GLCM calculation patch area. Large tissue areas under such low magnification such as this exhibit little regularity, and sample artefacts such as folding also disrupt pattern forming; therefore we did not predict high performance segmentation with this texture. Inflammation selected by one histopathologist in the top right of the tissue area is not demarcated on this figure as that portion of the tissue has severed and rotated away from its original position meaning a direct comparison here is not fruitful. Nonetheless, a comparison of this area, best selected in the energy (uniformity) texture map in Figure 64, with the hypothesised Raman inflammation segment (shown in blue in the aforementioned Figure) shows promising similarity. The homogeneity texture segmentation has performed remarkably well, ap-

proximately selecting unhealthy tissue (cancer and inflammation) and everything else (healthy tissue, holes and and cracks in the section) in two widespread segments. This is representative of the performance of this texture metric on all of the tongue SCC samples, shown in Figure 65.

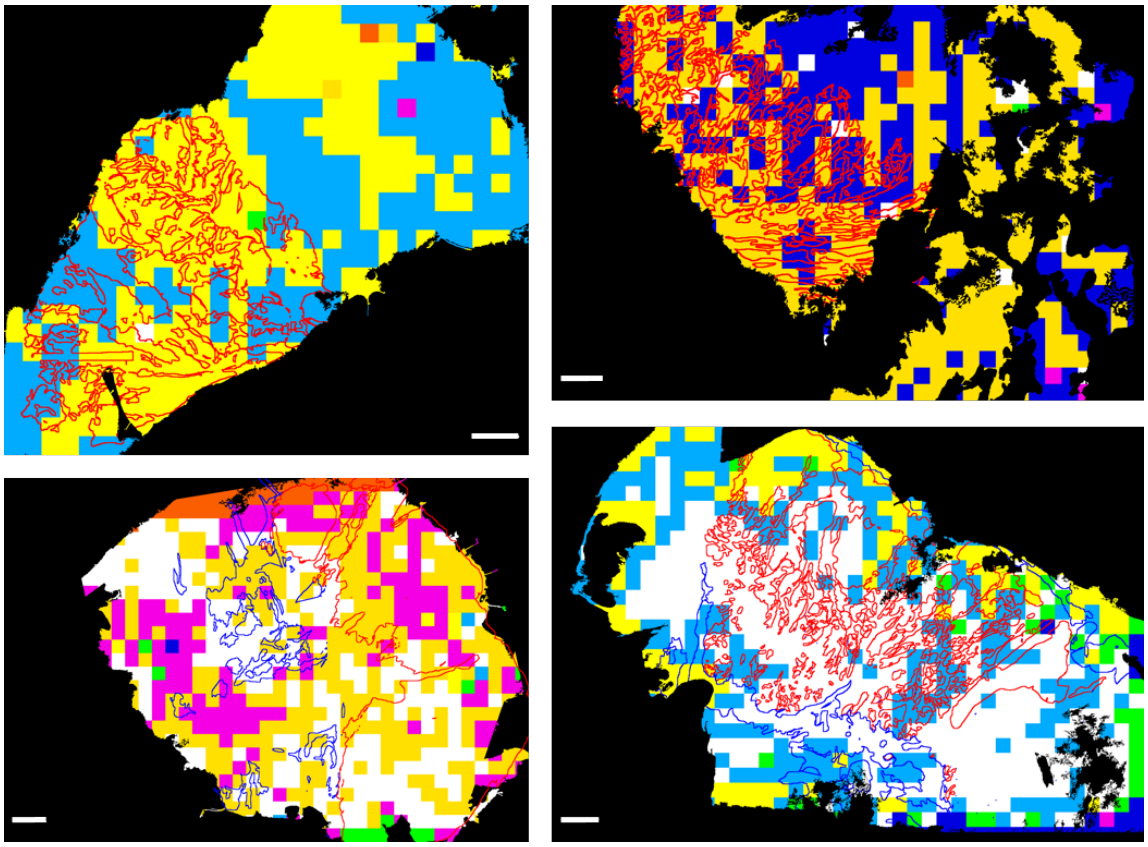

Figure 65: K-means clustered homogeneity texture-based segmentation of tongue squamous cell carcinoma samples #9806, #2577, #2217, #4644 (left to right, top to bottom). Consensus histopathology cancer segmentation is outlined in red. Inflammation demarcated by one histopathologist in royal blue on #2217 and #4644. Scale bar represents 500 µm.

Similarly to the inflammation point mentioned previously in relation to sample #9806, the perceived performance of the texture-based segmentation on these four samples is improved further with the consideration of the following points:

- Sample #2577 (top right): The histopathologists demarcated additional cancer areas on the fragmented section of tissue on the right, which supports the continuation of the orange cancer segment across the tissue break.
- Sample #2217 (bottom left): Two of the three histopathologists demarcated the main cancerous region extending further towards the centre of the tissue area, especially at the bottom which potentialy correlates better with the texture-based

segmentation, though the erroneous layer separation of this sample has somewhat affected classification performance by every method utilising that tissue section.

- Sample #4644: The histopathologist consensus demarcation ends cleanly at its left and right edges; this is due to a large hole in the H&E stained section on the right and separation and translation of the lobe on the left. These regions are therefore excluded from comparison.

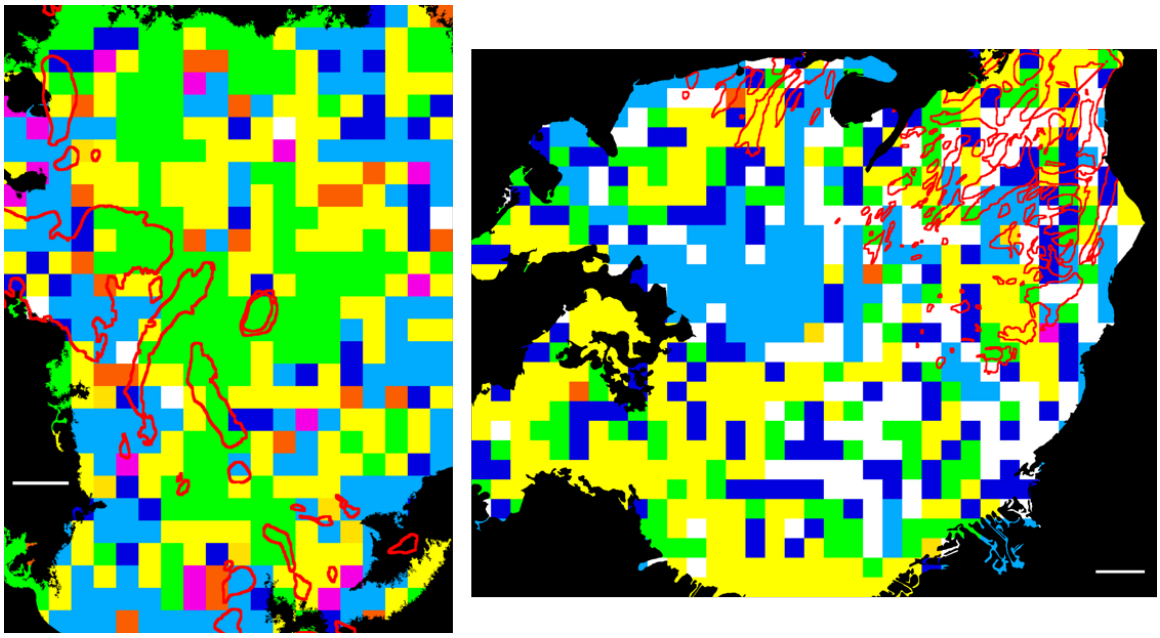

Figure 66: K-means clustered homogeneity texture-based segmentation of tongue squamous cell carcinoma samples #4734 and #9890 (left to right). Consensus histopathology cancer segmentation is outlined in red. Scale bar represents 500 µm.

As previously discussed, the random noise introduced by the tiling method depreciated the texture-based segmentation. Samples #4734 and #9890 (shown in Figure 66) , containing four field of view tiles rather than two, were worst affected again.

Once the parameters, such as patch size, are optimised and final improvements have been made to the imaging system and software this data should be re-collected to determine whether homogeneity maps are useful, and if texture segmentation could be improved by combining it with another texture metric with finer tissue clustering, such as contrast or energy. As ESS techniques are not chemically specific, healthy tissue is still classified into multiple segments, even though it only contains healthy tissue, as long as there are regions of textural difference, such as fibres.

While the texture based segmentation system still requires tweaking, it has other diagnostic value, in its current state: t-tests suggest that the contrast and energy standard deviations show statistically significant (p-value < 0.05) differences between the

healthy and SCC tongue tissue groups with p-values of 0.005 and 0.03 (1 s.f.), respectively, as shown in the box and whisker plot in Figure 67. Additionally, the difference between the mean contrast texture values of healthy tongue and tongue SCC is close to statistical significance with a p-value of 0.09. Drezek *et al* showed increased magnitude of forward scattering in cervical cells with increasing neoplastic progression.[107] Increased forward scattering would lead to increased pixel intensities with decreased uniformity, increasing contrast and energy texture values respectively.

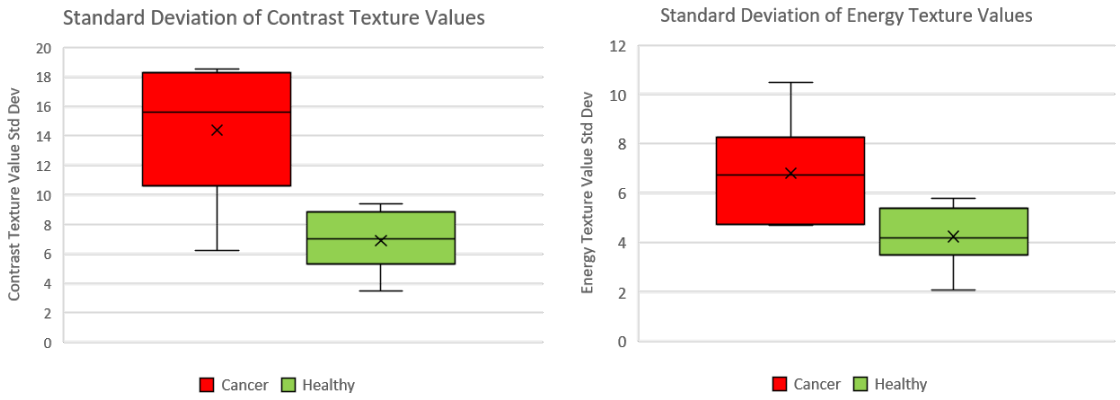

Figure 67: Box and whisker plots for contrast and energy texture value standard deviation,showing statistically significant differences between healthy tongue and tongue samples containing squamous cell carcinoma.

We can use this data set to create a diagnostic classifier based on one or both of these texture statistics. Testing should then be performed on a larger sample set of tongue squamous cell carcinomas, followed by other oral tissue sites and malignancies if it performs well in the first instance. Additionally, if the k-means clustering of the HSI data performs well after the suggested wavelength range truncation, the performance of this texture-based quick diagnostic screening would need to be re-evaluated with the new wavelength range. It could be a valuable time saving technique, used to exclude healthy samples from further investigation with Raman spectroscopy. Wavelength range truncation may lead to greater similarity between the contrast and energy standard deviations of cancerous and healthy oral tissue. Generally, but especially in this instance, multivariate analysis combining spatial (texture) features and spectral features may lead to the best possible diagnostic power and should be considered for future work.

A combination system in which the initial segmentation is performed by HSI, and the chemically specific/diagnostic info is provided by Raman sectroscopy would save a substantial amount of time compared to utilising Raman alone. This would allow for longer exposure times to be used for the Raman, if necessary, and especially if the

HSI-based segmentation could be trusted to the extent that only a handful of point-Raman-scans would be required to chemically evaluate each segment for diagnosis.

15.3 Additional Study - Bioaccumulation of Microplastics

15.3.1 Air Filter Samples

Air filter samples can contain a wide range of materials with 27% of total atmospheric fallout measured, in a study undertaken in China, confirmed to be plastic.[168] In a similar study in Paris, the range of microplastic particle diameters was 7 - 15 µm, suggesting that some may be present in this sample.[169]

As inhalable particles (smaller than 10 µm) were preferentially selected for this study due to their clinical relevance, the high spatial resolution of the HSI system was advantageous.[170] We would expect the transmission spectra of these small particles to show elements of absorption features (high frequency) and scattering features (low frequency). Depending on the relative dominance of each component, the segmentation capabilities of the HSI system could vary dramatically. This uncertainty is demonstrated in the segmentation map in Figure 68 below, produced by k-means clustering of the hypercube and manual selection and subsequent removal of the two background clusters: The cluster identities appear to be correlated with particle size which may be a result of scattering changes affected by particle packing density, where final particle packing density after drying of the droplet will be related to particle diameter and mass. It may be, however, that the clusters also correspond to different materials which are more likely to create particles of a certain size range.

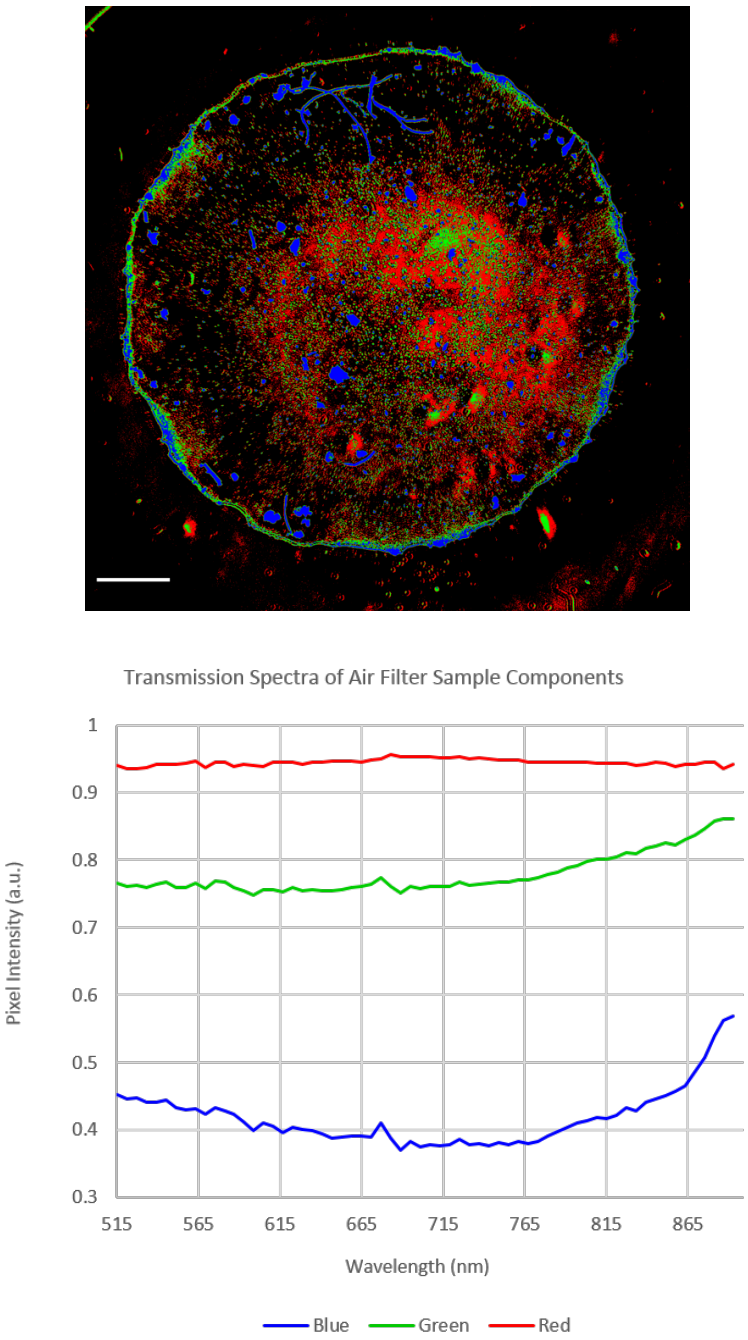

Figure 68: K-means clustered hypercube of dropcast of ambient particulate matter, reconstructed without the two background components for clarity (top). The mean transmission spectra of each of the three clusters, colour coded to match (bottom).

Raman is the ideal tool to ascertain the source of the spectral differences which produced this HSI clustering. Being a dropcast, coregistration between the HSI and Raman datasets would ordinarily be impossible, however in this instance the artefact at the top of the drop could be used as a marker. Preliminary Raman data was collected to probe the clustering basis of the HSI, and its precision by comparing the number and

location of particulate matter clusters identified in each modality with the knowledge that Raman segmentation is based on chemical difference. This data collection was not detailed in the Methods section (14.3.3) as only exploratory work has been possible so far. While this question remains unanswered, the Raman map showed dominating, distorted spectral responses correlating with the largest particles which do not align with a Raman response (separated into the blue cluster in Figure 69 below).

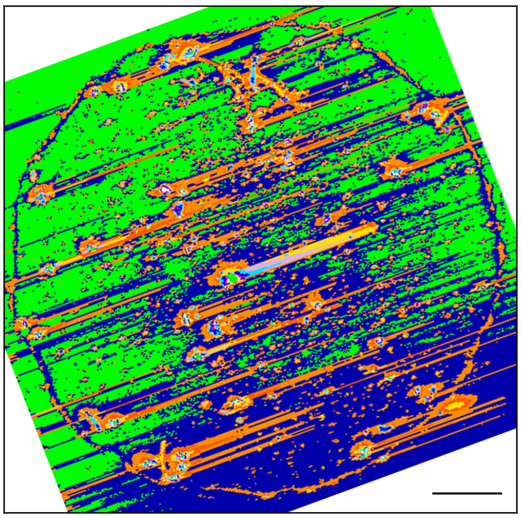

Figure 69: K-means clustered Raman map of ambient particulate matter dropcast showing suspected autofluorescence artefacts. The black scale bar represents 500 µm.

As laser reflection and Rayleigh scattering are filtered out inside the Raman InVia spectrometer, and resonant Mie scattering would afflict a smaller size range of particles, the current hypothesis is that autofluorescence is the source of these anomalies. This suggests that the blue channel represents biological material, or material with biological matter on its surface. This aligns with literature on air particulate matter samples which confirm that Raman spectral interference due to autofluorescence is common. One of the main confirmed causes of this is biofilms which grow on microplastic (and some other particle) surfaces, made possible by the microbial communities found in the near-atmosphere.[171] Alteration of the selected Raman wavenumber range for future work could diminish the impact of this autofluorescence.

15.3.2 Microplastic-Perfused Mouse Pulmonary Tissue

Scattering spectra extracted from within the bead areas of this hypercube, highlighted by the red boxes in Figure 70, show surprisingly large variation. The spherical shape of the microplastic inclusions in the pulmonary tissue, combined with the high spatial resolution of the system, means that scattering varies across the 2D image area of each bead. Different parts of the bead are therefore assigned to different clusters.

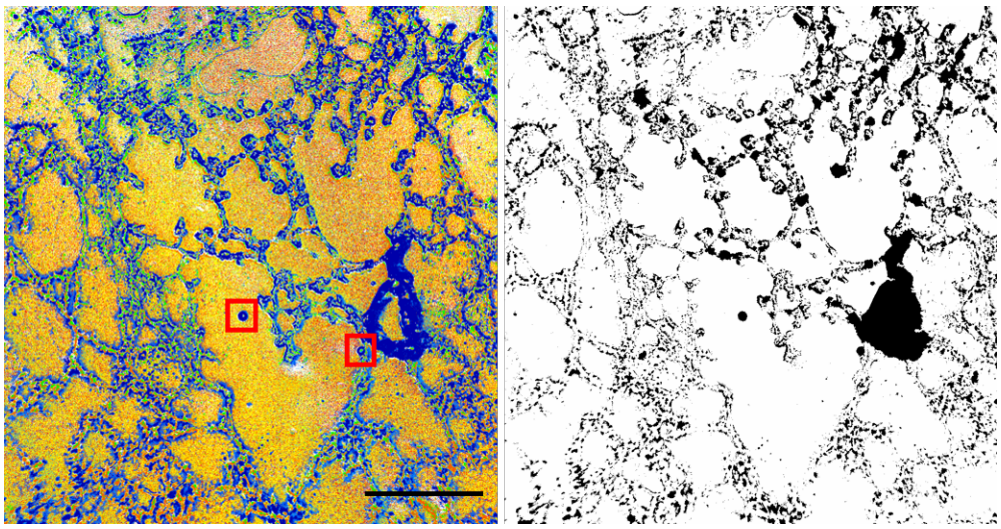

Figure 70: K-means clustered hypercube of mouse lung tissue perfused with colourless beads (left). A mask of the royal blue hue with filled holes (right).

This also happens, to some extent, with the tissue but the effect is less pronounced. As the tissue, perfused with beads up to 27 µm in diameter, is sectioned thinly to 10 µm; the contained beads are sectioned concurrently. Therefore, the depth-based cluster variations are assumed to be more apparent in the beads due to their three dimensionally symmetrical shape and consistent refractive index, where the refractive index variation of the pulmonary tissue is more inhomogeneous.

Decreasing the spatial resolution of the system by using a lower magnification objective may ameliorate these cluster "rings" by minimising detected spectral variation. However, the perfect circular cross section of the bead, and the aforementioned ring pattern of the clusters are the most distinctive features by which the bead could be identified. It is not yet clear whether the beads could be distinguished from the tissue without these features, at a lower magnification and resolution.

Plots of a range of spectra from each bead in this field of view (mean spectra shown in Figure 71) do show a smaller standard deviation than spectra from the tissue, with the two beads exhibiting mean spectral standard deviations of 0.47 and 0.85 (left to right) and the tissue, 1.08.

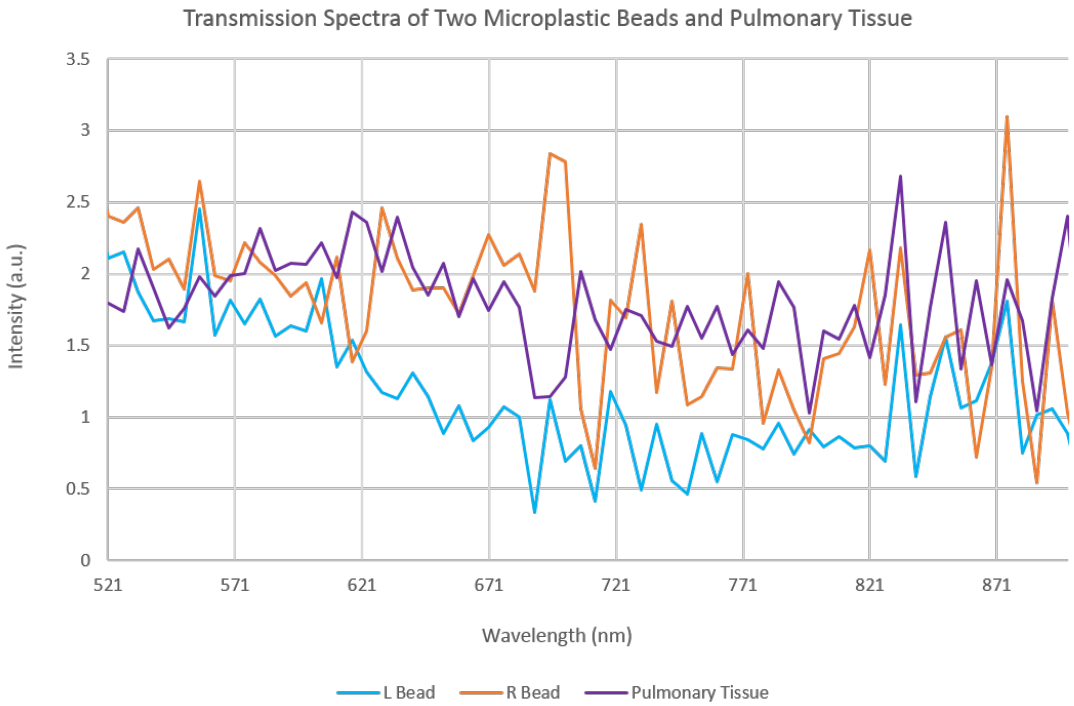

Figure 71: Graph of mean transmission spectra of the microplastic bead on the left of the field of view (L Bead) in Figure 70, the right hand bead (R Bead), and the pulmonary tissue itself.

This aligns with expectations as the tissue scattering variation is impacted by internal refractive index variation oweing to the distribution of subcellular components, whereas the bead scattering is impacted by a fixed range of depth variation due to section positioning. This increased standard deviation could potentially be used to discriminate the tissue from the beads. However, beads which are attached to, or immediately next to, the tissue of the alveolar sacs may not be successfully identified as separate scattering bodies requiring separate spectral standard deviation probing.

Texture analysis, or other spatial analysis, parameters could be constrained to detect circular particles which would be effective on this dataset (see the filled mask on the right of Figure 70). It is not, however, representative of environmental microplastics which would be the target sample in a screening protocol. In order to exploit a similar approach, any consistency in the shape and size of different environmental microplastic particles should be investigated.

The HSI spectra produced by all three dominant components in the particulate matter sample are sufficiently different to the spectra of colourless microplastic beads to facilitate their separation from each other: A mixed test sample will be created and imaged in confirmation of this assertion. The main questions posed by this preliminary

study relate to the separation of colourless microplastics from tissue. In tandem with investigating the size and shape distributons of environmental microplastics, their absorbant properties will be investigated in the hope of identifying a source of additional spectral contrast.

16 Future Work

Future work mostly pertains to the main study on the identification of oral cancer (in this study, specifically tongue squamous cell carcinoma) from unstained tissue sections; the primary motivation behind the development of this system.

Suggested future work can be broken into three categories; those which can be attempted with the existing data, those which require moderate alterations either to the hyperspectral microscope, the data collection software or the analysis software, and those which require substantial alterations to the system.

Firstly, the clinical viability of this potential HSI-Raman combination technique could be immediately bolstered by quantitatively evaluating the performance on the suggested wavelength subsets, 521 - 635 nm, 605 - 755 nm, which were extracted from the current hypercubes. (Their segmentation is shown in Figure 63.) Cutting down the wavelength range would decrease data collection time, and processing time. Qualitative evaluation shows promisingly similar segmentation in both ranges compared to the full wavelength range hypercube, though quantifying its effect on the sensitivity and specificity of the technique may impact decisions regarding future modifications to the system.

It would also be prudent to know some key wavelengths prior to creating diagnostic classifiers, which can begin in earnest now. While the k-means clustered maps of the wavelength-truncated HSI data appear very similar, the proposed coarse screening based on contrast and energy texture value standard deviations, for example, could be dramatically affected by wavelength cutting. Similarly, multivariate analysis which combines spatial texture metrics with spectral features as recommended in 15.2.5, may be greatly affected, although this analysis approach holds the greatest potential for high discriminatory power and should be investigated. Sensitivity and reliability must remain the priority through the process of identifying fixed diagnostic thresholds for the HSI pre-screening process. A wide range of classifiers have already been determined for Raman spectroscopy, prioritising specificity, which could be explored once a combination system is built.

Secondly, as combining spectral and spatial analysis approaches can improve overall classification performance and stabilise the "decision making" process of classification a spectral texture analysis programme has been written.[172] Spectral texture analysis utilises the same texture metrics calculated from the grey level co-occurrence matrix (GLCM) as spatial texture analysis. However, where the latter looks at each pixel and the kernel of pixels around it to calculate the texture for each wavelength, the former looks at each wavelength point on the spectrum and the kernel of spectral points around it to calculate the texture for each pixel. Providing similar information

to Fourier transform methods, but in the same texture value-encoded grayscale image stacks as the spatial texture outputs to facilitate simple combination of the spectral texture, spatial texture and raw HSI data or their k-means cluster maps in multivariate analysis. Following the exploration and analysis of the data in this work, parameters can be selected to begin work with this spectral texture analysis. It must be noted that the texture analysis further increases the dimensionality of the data which ultimately requires compressing into a segmented, 2D representation. With speed being an important feature of this proposed system, analysis timelines must be held in mind as additional analysis options are considered. The spatial texture analysis parameters used in this work also made it vulnerable to degradation by sample artefacts such as tears, folds and holes. While the homogeneity texture value appeared to contain some useful discriminatory information at this patch size (200 x 200 µm), a smaller patch size may be more robust when faced with tissue artefacts and should be tested.

In the data collection software, two changes could bring about a significant improvement in the SNR of the data: If the camera shutter could be changed from rolling to global, this would remove the horizontal striping artefact achieved for short exposure times and remove the need to perform maximum pixel acceptance corrections which are costly to the signal. This shutter settings change should be possible, but may require a change of platform for the data acquisition software. A new software platform may also facilitate the triggering of the pulsed Xenon source by the camera, which would result in a more consistent intensity at each wavelength in different datasets, helping prevent tiling artefacts. These are caused by the calculation of grayscale values in overlapping tile areas and introduce noise and segmentation errors at tile boundaries.

On a similar vein, the SNR and tiling errors could be further improved by utilising a lower magnification objective lense, such as x1.25 in place of x4, which would decrease the field of view tiling required to image one full tissue section. The increased angle of acceptance may also lead to a moderate increase in signal. One of the tongue squamous cell carcinoma samples from the main study, #4644, was imaged with a x1.25, 0.04 NA objective (Olympus) and k-means clustered with the parameters used in Figure15.2, but without coregistration (Figure 72).

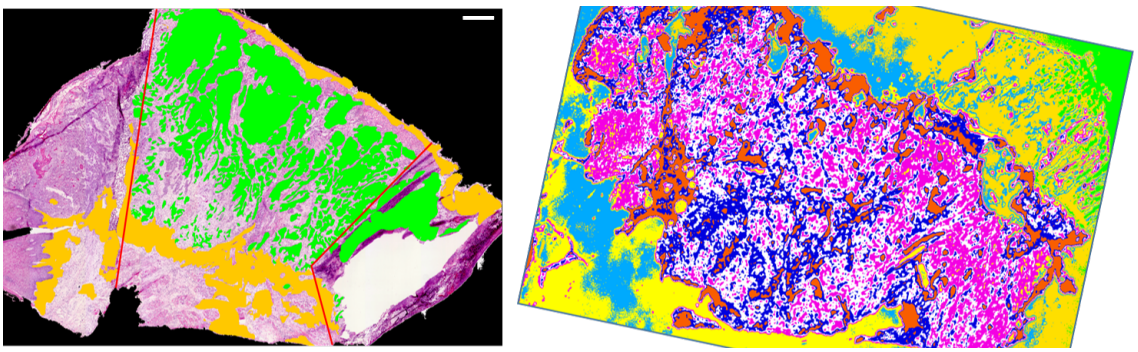

Figure 72: The RGB image of the H&E stained squamous cell carcinoma tongue tissue sample, #4644, with concensus histopathologist cancer demarcation in green, the inflammation demarcated by one histopathologist (orange) and areas of damaged or translated tissue areas separated by red lines, with only the central strip being considered for comparison (left). PCA dimensionality-reduced and k-means clustered HSI taken with a x1.25, 0.04 NA objective, with unmasked background (right).

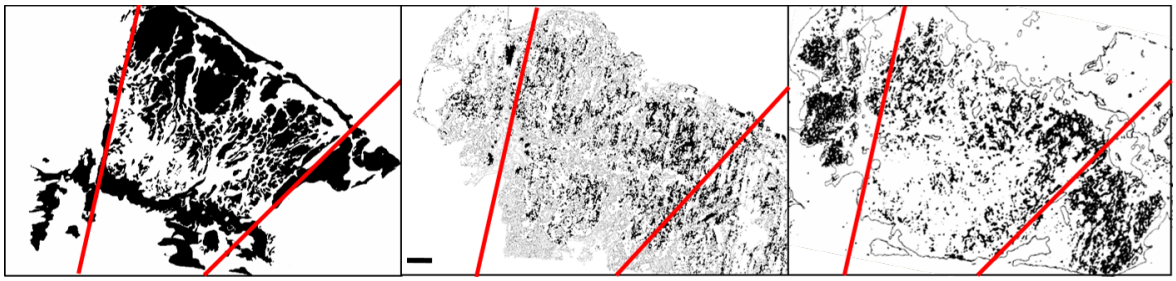

Figure 73: Masks of consensus histopathology (including inflammation), HSI x4 and HSI x1.25 cancer segments of tongue squamous cell carcinoma sample #4644 selected using the overlap coefficient macro with respect to histopathologist consensus diagnosis. Scale bar represents 500 µm.

Comparing the masks of the corresponding segments in the x1.25 and x4 segmented images shown in Figure 73 shows decreased sparsity with decreased magnification. It may be that a lower magnification and spatial resolution is more appropriate for making global maps of the tissue and its health status and that the x4 data was more sensitive to finer, perhaps subcellular, information which lead to less homogeneous region selection. For many reasons, this is worthy of further investigation.

Thirdly, and most significantly, more accurate inter-modal comparison could be achieved if we utilise the same frozen sections for all three imaging modes and if we build a HSI-Raman system.

Utilising the same frozen sections for histopathology, HSI and Raman would eliminate the most widespread sources of error across all tissue samples in this thesis; inherent differences between consecutive (or near-consecutive) tissue sections and dissimilar sampling artefacts between these sections. Achieving this is not a trivial problem: If we

aim to retrieve the histopathologists' diagnoses first, a H&E stain wash which does not leave a residual Raman spectrum will have to be identified. If we aim to stain it after collecting the HSI and Raman data, we would have to modify the method by which we sandwich it between two calcium fluoride slides to prevent drying out, shrinking and distorting over the long Raman collection process. It would have to remain airtight, but soft enough to allow separation of the slides for staining afterwards without being drawn into the sample by capillary action. Both of these considerations are surely possible, but may require a period of development.

Adding a HSI mode onto a Raman system would facilitate widefield HSI prescreening and subsequent Raman scanning for automated diagnosis with inherent coregistration. In addition to the HSI microscope components (12.1) our current Raman system (Renishaw InVia) would require only one major alteration; an automated objective lens turret to execute a switch between the HSI widefield objective (x4 or x1.25 magnification) to x20 for confocal Raman scanning. In order that a complex coregistration is not required, this objective turret must have high precision so that coordinates from the HSI segmentation can be directly applied, without affine transformation, to the Raman scan. Creating a hybrid system will also allow optimisation of the Raman data collection parameters which are chiefly limited by the time required to image a full tissue section with an acceptable resolution. The HSI-based pre-screening process will decrease the area which requires further probing with Raman. This will then permit use of a longer exposure time, increasing the SNR of the spectra and enhancing performance of segmentation software. A more accurate measure of performance can then be calculated to inform future work and feasibility of clinical translation.

17 Conclusion

This thesis outlined the development of a HSI-based cancer selection system for use with Raman spectroscopy for the detection of oral cancer from unstained, frozen tissue sections. Parts I and II explored the effects of various modifications to the system, including changes to the illumination source, and the development of a robust coregistration and evaluation workflow. With the final system and software in Part III, near-automation was achieved at every stage, with most user input required for the evaluation process rather than the data collection and segmentation process inherent to the system itself. Its performance was quantitatively compared to consensus histopathologist diagnosis, to evaluate sensitivity and specificity, and to Raman spectroscopy to confirm compatibility of the two techniques and the segmentation they produce.

This developmental work confirmed the importance of a powerful broadband white

light source with reliable intensity: A pulsed Xenon source is a suitable, economic choice and it is best suited to this application if triggering can be performed by the camera. Introduction of a reflective fibre collimator in Setup 2 dramatically ameliorated chromatic aberrations, improving focus across each field of view and removing the need for Z-stage control to account for a changing focal plane over the wavelength range. Use of diverging sample illumination prior to this setup addition, however, appeared to add a unique and useful scattering contribution. This may be explained by the acceptance of signal from higher scattering angles, potentially conveying sensitivity to more isotropic scattering such as Rayleigh scattering, over forward-biased Mie scattering. While this additional scattering appeared to contain useful information, the manifold benefits of using collimated light were prioritised given that adequate tissue information is retrieved without it.

Solutions to the two remaining limitations of the HSI microscope itself are detailed in the Future Work section (16): Use of a global camera shutter and camera-driven source pulse triggering would allow the removal of two data processing steps; taking the maximum intensity of each pixel over six dataset repeats, and the correction of small brightness variations across the wavelength range in spite of optimised exposure times. Exclusion of both of these steps would lead to improved SNR of the data and significantly decreased system run-time, improving classification and its suitability as a clinical diagnostic aid.

The near-total automation of the hyperspectral imaging, coregistration, segmentation and scoring means that the effects of modifications to any part of this process can be swiftly and easily quantified. Human input is only absolutely necessary in the selection of three sets of affine coordinates for the coregistration step (and for collecting the gold standard histopathologist' diagnosis for reference). Additionally, many of our software solutions are suitable for open-source distribution and were purposely developed on commonly used platforms; Matlab and ImageJ.

Aside from data collection equipment and processing methods, the largest source of error across all three Parts of this work come from the tissue sections themselves. Firstly, the inherent differences in tissue features between the consecutive (or near-consecutive) sections collected to facilitate the preparation of stained sections for histopathological review and unstained, frozen sections for collection of HSI and Raman spectroscopy data. Secondly, tissue sectioning artefacts such as breakage and folding which often occur dissimilarly between the two corresponding sections and complicate coregistration. This is unavoidable; the sectioning process is extremely delicate and these artefacts occurred in spite of careful sample preparation by an expert. Methods to minimise the impact of changing or erroneous tissue section features on the segmentation process

and its evaluation are detailed in Future Work (16). They include include development of a protocol which allows the same frozen section to be used for all modes, including histopathology, improving the accuracy of the performance evaluation. Subsequently, the building of a HSI-Raman combination system will remove the requirement for coregistration between the two modes, shedding more light on the compatibility of the two modes and producing diagnostic information in its final form.

The difficulties faced throughout this work relating to the sensitivity of Raman spectroscopy to sample storage and degradation would also be minimised in its intended clinical form and application. Histopathological review of tissue sections occurs in a timely manner after biopsy-collection and extended storage before initial evaluation would be extremely unlikely.

Histopathological segmentation was another source of uncertainty in the data: Using a qualitative technique as the gold standard is unusual outside of medical fields and carries inherent limitations. Interobserver variation was discussed and quantified, with a standard deviation of 6.7% across three histopathologists. To account for this substantial variance, the gold standard oral cancer segmentation was therefore defined as the consensus between all histopathologists, which produced conservative, but more reliable, cancer selections.

The low effective resolution of the preliminary texture analysis impeded evaluation using a ROC curve. However, statistically significant differences between healthy tongue tissue sections and tongue sections containing squamous cell carcinoma were observed: The difference in standard deviations of the contrast and energy texture values across hyperstacks of cancerous tissue and healthy tissue were both statistically significant ($p < 0.005$ and 0.03, respectively). This could be exploited for an initial exclusion protocol, excluding healthy samples from further examination by histopathologists or the HSI-Raman system: This could be a significant time-saving step, well suited to improving oral cancer survival rates which suffer from late detection. Differences between mean contrast texture values in tongue squamous cell carcinoma and healthy tongue were close to statistical significance ($p < 0.09$), and would likely become significant after the proposed modifications to the system and their associated SNR improvement.

PCA denoising and k-means clustering of the HSI data provided the best segmentation, and the quickest analysis time ($<$ 30 minutes for one field of view), garnering an AUROC score of 0.70 with respect to the gold standard. Comparison with the two most similar studies rate its performance very differently. Fei *et al* achieved a diagnostic sensitivity, and specificity of 66% and 76% in ex-vivo tumour-normal interface tissue. Given the presence of intensely absorbant and clinically informative haemoglobin in

their samples, which are minimal in our unstained thin tissue sections, our system has performed well and would not require remodelling of the diagnostic procedure for widespread use. With this system they were able to garner an AUROC score of 0.88 for sample-wise diagnosis. Given our similar sensitivity and specificity (dependent on selection of diagnostic threshold), we might hope to achieve a similar AUROC with our current system for sample-wise diagnosis.[155] On the other hand, performance is poor compared to the 96% sensitivity and 100% specificity achieved with the quantum cascade laser-based IR microspectrometer system from the work of Kuepper *et al.*[156] However, this work is performed on colorectal cancer which may have less intraclass variation and the system has a lower spatial resolution than Raman spectroscopy and contains several expensive components which may prohibit its clinical translation.

It may be the case that, once the suggested modifications are made to the system and diagnostic classifiers are identified from the k-means segmentation already achieved, the HSI pre-screening process will be sufficient as a standalone oral cancer detection system. As it stands, however, the HSI system does not reliably separate inflammation and cancer and achieved an AUROC score of 0.70 with respect to histopathology. Adding the chemical specificity of Raman spectroscopy by creating a hybrid system would ensure successful separation and the AUROC score of the HSI segmentation technique with respect to Raman (0.69) suggests that their segmentation is compatible. This should be re-evaluated once the system modifications are complete, and again once the hybrid HSI-Raman system is constructed. Diagnostic performance of the final, hybrid system with AUROC score of > 0.90 with respect to the gold standard histopathology, and with an AUROC score >0.85 when evaluating the HSI pre-screening segmentation with respect to the Raman (given that areas of inflammation are likely to be in disagreement between the HSI and Raman modes) would warrant pursuit of clinical translation. This would be especially worthwhile if the suggested wavelength truncation for speed is successful. Given the range and magnitude of limitations in the current system, this may be achievable.

Preliminary work suggests that the system may be suitable for application in the screening for bioaccumulated particulate matter and microplastics, which are a growing source of interest and concern. Further work with environmental plastic samples is required to confirm this, with suitable samples being readily available.

A HSI pre-screening system was constructed which (if the camera shutter mode can be altered from rolling to global, reducing collection time six-fold) can collect a hypercube of a tissue section up to 4 x 4 mm in area across a 521 - 899 nm wavelength

range in ~ 15 minutes with a spectral resolution of 6 nm and spatial resolution of ~ 2 μm. Hyperstacks across substantially smaller wavelength ranges (521 - 635 nm or 605 - 755 nm) appear to produce similar segmentation, though performance has not been quantitatively evaluated. Affine transformation-based coregistration followed by PCA denoising and k-means clustering appears to produce the most diagnostically relevant clustering, garnering a pixel-wise cancer segmentation performance AUROC score of 0.70 with respect to consensus histopathologist diagnosis in $<$ 30 minutes per field of view. Although texture analysis parameters have yet to be fully explored, preliminary work shows statistically significant differences between cancer and non-cancer in the standard deviations of contrast ($p < 0.005$) and energy ($p < 0.03$) texture values. The segmentation provided by Raman spectroscopy correlates well with it, and a hybrid HSI-Raman system is proposed to exploit the speed of the former and the specificity and proven diagnostic power of the latter. The HSI system alone, or with Raman spectroscopy, has the potential to fit into the current diagnostic flow with no additional preparation or prior knowledge of the sample required, to support histopathologists in the quest for early diagnosis of oral cancer to improve survival rates.

18 References

[1] David I Conway, Mitana Purkayastha, and Ivor G Chestnutt. The changing epidemiology of oral cancer: definitions, trends, and risk factors. *British dental journal*, 225(9):867–873, 2018.

[2] Oral Health Foundation. State of Mouth Cancer UK Report 2020/21. 2020.

[3] Saman Warnakulasuriya. Global epidemiology of oral and oropharyngeal cancer. *Oral oncology*, 45(4-5):309–316, 2009.

[4] Stefano Tiziani, Victor Lopes, and Ulrich L Günther. Early stage diagnosis of oral cancer using 1h nmr–based metabolomics. *Neoplasia*, 11(3):269–IN10, 2009.

[5] John D McDowell. An overview of epidemiology and common risk factors for oral squamous cell carcinoma. *Otolaryngologic Clinics of North America*, 39(2):277–294, 2006.

[6] Raymond J Melrose, Janice P Handlers, Stanley Kerpel, Don-John Summerlin, Charles J Tomich, et al. The use of biopsy in dental practice. the position of the american academy of oral and maxillofacial pathology. *General dentistry*, 55(5):457–61, 2007.

[7] Simon N Rogers, R Pabla, A McSorley, Derek Lowe, James S Brown, and E David Vaughan. An assessment of deprivation as a factor in the delays in presentation, diagnosis and treatment in patients with oral and oropharyngeal squamous cell carcinoma. *Oral oncology*, 43(7):648–655, 2007.

[8] Juan Seoane, Pablo Varela-Centelles, Germán Esparza-Gómez, Rocío Cerero-Lapiedra, Juan M Seoane-Romero, and Pedro Diz. Simulation for training in oral cancer biopsy: A surgical model and feedback from gdps. *Medicina oral, patologia oral y cirugia bucal*, 18(2):e246, 2013.

[9] Roderick A Cawson and Edward W Odell. *Cawson's essentials of oral pathology and oral medicine*. Churchill Livingstone, 2008.

[10] Wusheng Yan, Ignacio I Wistuba, Michael R Emmert-Buck, and Heidi S Erickson. Squamous Cell Carcinoma - Similarities and Differences among Anatomical Sites. *American journal of cancer research*, 1(3):275–300, 2011.

[11] Zachary S Peacock, M Anthony Pogrel, and Brian L Schmidt. Exploring the reasons for delay in treatment of oral cancer. *Journal of the American Dental Association (1939)*, 139(10):1346–1352, 2008.

[12] Jia Hui Ng, N Gopalakrishna Iyer, Min-Han Tan, and Gustaf Edgren. Changing epidemiology of oral squamous cell carcinoma of the tongue: A global study. *Head & neck*, 39(2):297–304, 2017.

[13] NHS HPV Human Papillomavirus Vaccine. https://www.nhs.uk/conditions/vaccinations/hpv-human-papillomavirus-vaccine . Accessed on 29-05-21.

[14] Juliana Noguti, Carolina Foot Gomes De Moura, Gustavo Protasio Pacheco De Jesus, Victor Hugo Pereira Da Silva, Thais Ayako Hossaka, Celina Tijuko Fujiyama Oshima, and Daniel Araki Ribeiro. Metastasis from oral cancer: an overview. *Cancer genomics & proteomics*, 9(5):329–35, 2012.

[15] Alfio Ferlito, Ashok R Shaha, Carl E Silver, Alessandra Rinaldo, and Vanni Mondin. Incidence and sites of distant metastases from head and neck cancer. *ORL*, 63(4):202–207, 2001.

[16] Jane Lea, Gideon Bachar, Anna M Sawka, Deepak C Lakra, Ralph W Gilbert, Jonathan C Irish, Dale H Brown, Patrick J Gullane, and David P Goldstein. Metastases to Level IIb in Squamous Cell Carcinoma of the Oral Cavity: A Systematic Review and Meta-Analysis. *Head and Neck - Journal for the Sciences and Specialties of the Head and Neck*, 32(2):184–190, 2010.

[17] Song Fan, Qiong-Lan Tang, Ying-Jin Lin, Wei-Liang Chen, Jin-Song Li, Zhi-Quan Huang, Zhao-Hui Yang, You-Yuan Wang, Da-Ming Zhang, Hui-Jing Wang, Eduardo Dias-Ribeiro, Qiang Cai, and Lei Wang. A review of clinical and histological parameters associated with contralateral neck metastases in oral squamous cell carcinoma. *Int J Oral Sci International Journal of Oral Science*, 3(3):180–191, 2011.

[18] Jayson S. Greenberg, Robert Fowler, Jose Gomez, Vivian Mo, Dianna Roberts, Adel K. El Naggar, and Jeffrey N. Myers. Extent of extracapsular spread: A critical prognosticator in oral tongue cancer. *Cancer*, 97(6):1464–1470, 2003.

[19] Sridhar Ramaswamy, Ken N Ross, Eric S Lander, and Todd R Golub. A molecular signature of metastasis in primary solid tumors. *Nature genetics*, 33(1):49–54, 2003.

[20] Gordon B Snow, Michiel W M van den Brekel, C Renee Leemans, and P. Patel. Surgical Management of Cervical Lymph Nodes in Patients with Oral and Oropharyngeal Cancer. pages 43–55. Springer Berlin Heidelberg, 1994.

[21] Steven Woeste. Second opinions for pathology reports? *Laboratory Medicine*, 32(11):665–669, 2001.

[22] Andreas Staude and Jürgen Goebbels. Determining the spatial resolution in computed tomography-comparison of mtf and line-pair structures. 2011.

[23] Jeffrey M Albert. Radiation risk from ct: implications for cancer screening. *American Journal of Roentgenology*, 201(1):W81–W87, 2013.

[24] Bingsheng Huang, Martin Wai-Ming Law, and Pek-Lan Khong. Whole-body pet/ct scanning: estimation of radiation dose and cancer risk. *Radiology*, 251(1):166–174, 2009.

[25] Donald G Mitchell, David Lawrence Burk Jr, Simon Vinitski, and Matthew D Rifkin. The biophysical basis of tissue contrast in extracranial mr imaging. *American Journal of Roentgenology*, 149(4):831–837, 1987.

[26] Eric C Ehman, Geoffrey B Johnson, Javier E Villanueva-Meyer, Soonmee Cha, Andrew Palmera Leynes, Peder Eric Zufall Larson, and Thomas A Hope. Pet mri: where might it replace pet ct? *Journal of Magnetic Resonance Imaging*, 46(5):1247–1262, 2017.

[27] Roopa S Rao, Kasimsetty Ramakantha Chatura, SV Sowmya, Kavitha Prasad, Surendra Lakshminarayana, Fareedi Mukram Ali, Kamran Habib Awan, and Shankargouda Patil. Procedures and pitfalls in incisional biopsies of oral squamous cell carcinoma with respect to histopathological diagnosis. *Disease-a-Month*, page 101035, 2020.

[28] Sara Chen, Michael Forman, Peter M Sadow, and Meredith August. The diagnostic accuracy of incisional biopsy in the oral cavity. *Journal of Oral and Maxillofacial Surgery*, 74(5):959–964, 2016.

[29] Priyanka Ramesh, Ramanathan Karuppasamy, and Shanthi Veerappapillai. A review on recent advancements in diagnosis and classification of cancers using artificial intelligence. *BioMedicine*, 10(3):5, 2020.

[30] Tony Wilson. Resolution and optical sectioning in the confocal microscope. *Journal of microscopy*, 244(2):113–121, 2011.

[31] Sezgin M Ismail, Angela B Colclough, John S Dinnen, Douglas Eakins, David M Evans, Ernest Gradwell, Jerry P O'Sullivan, Joan M Summerell, and Robert G

Newcombe. Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia. *BMJ (Clinical research ed.)*, 298(6675):707–10, mar 1989.

[32] W Glenn McCluggage, Hoshang Bharucha, Linda M Caughley, A Date, Peter W Hamilton, Colleen M Thornton, and Maureen Y Walsh. Interobserver variation in the reporting of cervical colposcopic biopsy specimens: comparison of grading systems. *Journal of clinical pathology*, 49(10):833–5, 1996.

[33] David M Melville, Jeremy R Jass, Neil A Shepherd, J M Northover, D Capellaro, P I Richman, J E Lennard-Jones, J K Ritchie, and S N Andersen. Dysplasia and deoxyribonucleic acid aneuploidy in the assessment of precancerous changes in chronic ulcerative colitis. Observer variation and correlations. *Gastroenterology*, 95(3):668–75, sep 1988.

[34] Cook Deborah Odell Ian D. Immunofluorescence Techniques. *Journal of Investigative Dermatology*, 133, 2013.

[35] Clifford B Saper. A guide to the perplexed on the specificity of antibodies. *Journal of Histochemistry & Cytochemistry*, 57(1):1–5, 2009.

[36] Jennifer Sturgis. | immunofluorescence. *Education Guide Immunohistochemical Staining Methods Fifth Edition*, page 61, 2009.

[37] Ravikant S Ganga, Dipali Gundre, Shivani Bansal, Pankaj M Shirsat, Pooja Prasad, and Rajiv S Desai. Evaluation of the diagnostic efficacy and spectrum of autofluorescence of benign, dysplastic and malignant lesions of the oral cavity using velscope. *Oral oncology*, 75:67–74, 2017.

[38] Daniel Navarini, Richard R Gurski, Carlos Augusto Madalosso, Lucas Aita, Luise Meurer, and Fernando Fornari. Epidermal growth factor receptor expression in esophageal adenocarcinoma: relationship with tumor stage and survival after esophagectomy. *Gastroenterology research and practice*, 2012, 2012.

[39] Priyanka Sawant, Avinash Kshar, Raghvendra Byakodi, and Arati Paranjpe. Immunofluorescence in Oral Mucosal Diseases - A Review. *Oral Surgery, Oral Medicine, Oral Radiology*, 2(1):6–10, 2014.

[40] Sandhya Avinash Tamgadge, Sindhu Milind Ganvir, Vinay Krishnarao Hazarey, and Avinash Tamgadge. Oral leukoplakia: Transmission electron microscopic correlation with clinical types and light microscopy. *Dental research journal*, 9(Suppl 1):S94–S104, dec 2012.

[41] Michael Brennan and Migliorati et al. Management of oral epithelial dysplasia: a review. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 103:S19.e1–S19.e12, mar 2007.

[42] Richard C K Jordan, Maricris Macabeo-Ong, Caroline H Shiboski, Nusi Dekker, David G Ginzinger, David T W Wong, and Brian L Schmidt. Overexpression of matrix metalloproteinase-1 and -9 mRNA is associated with progression of oral dysplasia to cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 10(19):6460–5, oct 2004.

[43] Sanatombi Rajkumari, Jeevakarunyam Sathiyajeeva, Chitta Suresh Kumar, Paramel Mohan Sunil, and Balakrishnan Thayumanavan. Molecular predictors in the early diagnosis of oral cancer. *Journal of clinical and diagnostic research : JCDR*, 7(5):942–4, may 2013.

[44] Petra Wilder-Smith, Woong-Gyu Jung, Matthew Brenner, Kathryn Osann, Hamza Beydoun, Diana Messadi, and Zhongping Chen. In vivo optical coherence tomography for the diagnosis of oral malignancy. *Lasers in Surgery and Medicine: The Official Journal of the American Society for Laser Medicine and Surgery*, 35(4):269–275, 2004.

[45] Elizabeth J Franzmann, Erika P Reategui, Kermit L Carraway, Kara L Hamilton, Donald T Weed, and W Jarrard Goodwin. Salivary soluble cd44: a potential molecular marker for head and neck cancer. *Cancer Epidemiology and Prevention Biomarkers*, 14(3):735–739, 2005.

[46] Chih-Yu Wang, Tsuimin Tsai, Hui-Chun Chen, Shu-Chen Chang, Chin-Tin Chen, and Chun-Pin Chiang. Autofluorescence spectroscopy for in vivo diagnosis of dmba-induced hamster buccal pouch pre-cancers and cancers. *Journal of oral pathology & medicine*, 32(1):18–24, 2003.

[47] Wei Zheng, Khee Chee Soo, Ranjiv Sivanandan, and Malini Olivo. Detection of neoplasms in the oral cavity by digitized endoscopic imaging of 5-aminolevulinic acid-induced protoporphyrin ix fluorescence. *International journal of oncology*, 21(4):763–768, 2002.

[48] Kum Ju Chae, Hyunsook Hong, Soon Ho Yoon, Seokyung Hahn, Gong Yong Jin, Chang Min Park, and Jin Mo Goo. non-diagnostic results of percutaneous transthoracic needle biopsy: A meta-analysis. *Scientific reports*, 9(1):1–12, 2019.

[49] Shlomo Efrima and Horia Metiu. Classical theory of light scattering by a molecule located near a solid surface. *Chemical Physics Letters*, 60(1):59–64, 1978.

[50] Andrew Downes. Raman microscopy and associated techniques for label-free imaging of cancer tissue. *Applied Spectroscopy Reviews*, 50(8):641–653, 2015.

[51] Catherine Kendall, Joanne Hutchings, Hugh Barr, Neil Shepherd, and Nicholas Stone. Exploiting the diagnostic potential of biomolecular fingerprinting with vibrational spectroscopy. *Faraday discussions*, 149(1):279–290, 2011.

[52] Gerwin J Puppels, FFM De Mul, Cees Otto, Jan Greve, Michel Robert-Nicoud, Donna J Arndt-Jovin, and Thomas M Jovin. Studying single living cells and chromosomes by confocal raman microspectroscopy. *Nature*, 347(6290):301–303, 1990.

[53] Monika Gniadecka, HC Wulf, N Nymark Mortensen, O Faurskov Nielsen, and DH Christensen. Diagnosis of basal cell carcinoma by raman spectroscopy. *Journal of Raman Spectroscopy*, 28(2-3):125–129, 1997.

[54] L Su, YF Sun, Y Chen, P Chen, AG Shen, XH Wang, J Jia, YF Zhao, XD Zhou, and JM Hu. Raman spectral properties of squamous cell carcinoma of oral tissues and cells. *Laser Physics*, 22(1):311–316, 2012.

[55] Po-Hsiung Chen, Rintaro Shimada, Sohshi Yabumoto, Hajime Okajima, Masahiro Ando, Chiou-Tzu Chang, Li-Tzu Lee, Yong-Kie Wong, Arthur Chiou, and Hiro-o Hamaguchi. Automatic and objective oral cancer diagnosis by raman spectroscopic detection of keratin with multivariate curve resolution analysis. *Scientific reports*, 6(1):1–9, 2016.

[56] Elisa M Barroso, Roeland WH Smits, Cornelia GF van Lanschot, Peter J Caspers, Ivo Ten Hove, Hetty Mast, Aniel Sewnaik, José A Hardillo, Cees A Meeuwis, Rob Verdijk, et al. Water concentration analysis by raman spectroscopy to determine the location of the tumor border in oral cancer surgery. *Cancer research*, 76(20):5945–5953, 2016.

[57] Isha Behl, Lekha Kukreja, Atul Deshmukh, SP Singh, Hitesh Mamgain, Arti R Hole, and C Murali Krishna. Raman mapping of oral buccal mucosa: a spectral histopathology approach. *Journal of biomedical optics*, 19(12):126005, 2014.

[58] Luis Felipe CS Carvalho, Franck Bonnier, Kate O'Callaghan, Jeff O'Sullivan, Stephen Flint, Hugh J Byrne, and Fiona M Lyng. Raman micro-spectroscopy for rapid screening of oral squamous cell carcinoma. *Experimental and molecular pathology*, 98(3):502–509, 2015.

[59] Chi Zhang, Delong Zhang, and Ji-Xin Cheng. Coherent raman scattering microscopy in biology and medicine. *Annual review of biomedical engineering*, 17:415–445, 2015.

[60] Halina Abramczyk. *Introduction to laser spectroscopy*. Elsevier, 2005.

[61] Christoph Krafft, Benjamin Dietzek, Jürgen Popp, and Michael Schmitt. Raman and coherent anti-stokes raman scattering microspectroscopy for biomedical applications. *Journal of biomedical optics*, 17(4):040801, 2012.

[62] Nada N Boustany, Stephen A Boppart, and Vadim Backman. Microscopic imaging and spectroscopy with scattered light. *Annual review of biomedical engineering*, 12:285–314, 2010.

[63] Richard O Prum, Rodolfo H Torres, Scott Williamson, and Jan Dyck. Coherent light scattering by blue feather barbs. *Nature*, 396(6706):28–29, 1998.

[64] Craig F Bohren and Donald R Huffman. *Absorption and scattering of light by small particles*. John Wiley & Sons, 2008.

[65] Toru Yoshizawa and Chapter 12 by Lev Perelman. *Handbook of optical metrology: principles and applications*. CRC press, 2017.

[66] Judith R Mourant, James P Freyer, Andreas H Hielscher, Angelia A Eick, Dan Shen, and Tamara M Johnson. Mechanisms of light scattering from biological cells relevant to noninvasive optical-tissue diagnostics. *Applied optics*, 37(16):3586–3593, 1998.

[67] JM Schmitt and G Kumar. Turbulent nature of refractive-index variations in biological tissue. *Optics letters*, 21(16):1310–1312, 1996.

[68] DA Zimnyakov, VV Tuchin, and AA Mishin. Spatial speckle correlometry in applications to tissue structure monitoring. *Applied Optics*, 36(22):5594–5607, 1997.

[69] Lev T Perelman, Mark D Modell, Edward Vitkin, and Eugene B Hanlon. Light scattering spectroscopy: from elastic to inelastic. *Handbook of Coherent-Domain Optical Methods*, page 437, 2013.

[70] Hendrik Christoffel van de Hulst. *Light scattering by small particles*. Courier Corporation, 1981.

[71] Zachary A Steelman, Derek S Ho, Kengyeh K Chu, and Adam Wax. Light-scattering methods for tissue diagnosis. *Optica*, 6(4):479–489, 2019.

[72] Paul Bassan, Hugh J Byrne, Franck Bonnier, Joe Lee, Paul Dumas, and Peter Gardner. Resonant mie scattering in infrared spectroscopy of biological materials–understanding the dispersion artefact. *Analyst*, 134(8):1586–1593, 2009.

[73] R de L Kronig. On the theory of dispersion of x-rays. *Josa*, 12(6):547–557, 1926.

[74] Condon Lau, Obrad R Scepanovic, Jelena Mirkovic, Sasha McGee, Chung-Chieh Yu, Stephen F Fulghum Jr, Michael B Wallace, James W Tunnell, Kate L Bechtel, and Michael S Feld. Re-evaluation of model-based light-scattering spectroscopy for tissue spectroscopy. *Journal of biomedical optics*, 14(2):024031, 2009.

[75] Paul Bassan, Achim Kohler, Harald Martens, Joe Lee, Hugh J Byrne, Paul Dumas, Ehsan Gazi, Michael Brown, Noel Clarke, and Peter Gardner. Resonant mie scattering (rmies) correction of infrared spectra from highly scattering biological samples. *Analyst*, 135(2):268–277, 2010.

[76] Roberto Reif, Mark S Amorosino, Katherine W Calabro, Ousama M Aamar, Satish K Singh, and Irving J Bigio. Analysis of changes in reflectance measurements on biological tissues subjected to different probe pressures. *Journal of biomedical optics*, 13(1):010502, 2008.

[77] A Sharwani, W Jerjes, V Salih, B Swinson, IJ Bigio, M El-Maaytah, and C Hopper. Assessment of oral premalignancy using elastic scattering spectroscopy. *Oral oncology*, 42(4):343–349, 2006.

[78] Susan G Brouwer de Koning, Elisabeth JM Baltussen, M Baris Karakullukcu, Behdad Dashtbozorg, Laura A Smit, Richard Dirven, Benno HW Hendriks, Henricus JCM Sterenborg, and Theo JM Ruers. Toward complete oral cavity cancer resection using a handheld diffuse reflectance spectroscopy probe. *Journal of biomedical optics*, 23(12):121611, 2018.

[79] Rupananda Jayachandra Mallia, Shiny Sara Thomas, Anitha Mathews, Rejnish Kumar, Paul Sebastian, Jayaprakash Madhavan, and Narayanan Subhash. Oxygenated hemoglobin diffuse reflectance ratio for in vivo detection of oral precancer. *Journal of biomedical optics*, 13(4):041306, 2008.

[80] Susan G Brouwer de Koning, MB Karakullukcu, L Smit, EJM Baltussen, HJCM Sterenborg, and TJM Ruers. Near infrared hyperspectral imaging to evaluate tongue tumor resection margins intraoperatively (conference presentation). In *Optical Imaging, Therapeutics, and Advanced Technology in Head and Neck Surgery and Otolaryngology 2018*, volume 10469, page 104690G. International Society for Optics and Photonics, 2018.

[81] Spectral Imaging Ltd. John R. Gilchrist, Gilden Photonics Ltd.; Timo Hyvärinen. Hyperspectral Imaging Spectroscopy, 2002.

[82] Jasper G Tallada, Pepito M Bato, Bim P Shrestha, Taichi Kobayashi, and Masateru Nagata. Quality evaluation of plant products. In *Hyperspectral Imaging Technology in Food and Agriculture*, pages 227–249. Springer, 2015.

[83] Rami Saari, Heikki and Antila, Jarkko and Holmlund, Christer and Mannila and Kai Mäkynen, Jussi and Viherkanto. Fabry-Perot Interferometer Hyperspectral Imaging Technology Transfer to Space Applications 8th ESA Round Table on Micro and Nano Technologies for Space Applications Session 6 : Optical Sensors & Actuators. In *8th ESA Round Table on Micro and Nano Technologies for Space Applications*, 2012.

[84] Sascha Grusche. Basic slit spectroscope reveals three-dimensional scenes through diagonal slices of hyperspectral cubes. *Applied optics*, 53(20):4594–603, 2014.

[85] Evencio Mediavilla, Santiago Arribas, Martin Roth, Jordi Cepa-Nogué, S. F Sánchez, and . 3D Spectroscopy in Astronomy. *Igarss 2014*, (1):1–5, 2014.

[86] Andrew Bodkin, A Sheinis, A Norton, J Daly, S Beaven, and J Weinheimer. Snapshot hyperspectral imaging: the hyperpixel array camera. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV*, volume 7334, page 73340H. International Society for Optics and Photonics, 2009.

[87] A. A. Gowen, C. P. O'Donnell, P. J. Cullen, G. Downey, and J. M. Frias. Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trends in Food Science and Technology*, 18(12):590–598, 2007.

[88] Vincent J. Realmuto. Summaries of the Fifth Annual JPL Airborne Earth Science Workshop. Volume 2: TIMS Workshop. 1995.

[89] Yasser Khouj, Jeremy Dawson, James Coad, and Linda Vona-Davis. Hyperspectral imaging and k-means classification for histologic evaluation of ductal carcinoma in situ. *Frontiers in oncology*, 8:17, 2018.

[90] Weiwei Ma, Cailan Gong, Yong Hu, Peng Meng, and Feifei Xu. The hughes phenomenon in hyperspectral classification based on the ground spectrum of grasslands in the region around qinghai lake. In *International Symposium on Photoelectronic Detection and Imaging 2013: Imaging Spectrometer Technologies and Applications*, volume 8910, page 89101G. International Society for Optics and Photonics, 2013.

[91] Hamed Akbari, Yukio Kosugi, Kazuyuki Kojima, and Naofumi Tanaka. Blood vessel detection and artery-vein differentiation using hyperspectral imaging. *31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1461–1464, 2009.

[92] Hamed Akbari, Kuniaki Uto, Yukio Kosugi, Kazuyuki Kojima, and Naofumi Tanaka. Cancer detection using infrared hyperspectral imaging. *Cancer Science*, 102(4):852–857, 2011.

[93] Hamed Akbari, Luma V Halig, David M Schuster, Adeboye Osunkoya, Viraj Master, Peter T Nieh, Georgia Z Chen, and Baowei Fei. Hyperspectral imaging and quantitative analysis for prostate cancer detection. *Journal of biomedical optics*, 17(7):076005, 2012.

[94] Hamed Akbari, Luma V Halig, Hongzheng Zhang, Dongsheng Wang, Zhuo Georgia Chen, and Baowei Fei. Detection of Cancer Metastasis Using a Novel Macroscopic Hyperspectral Method. *Proceedings of SPIE*, 8317:831711, 2012.

[95] Seong G. Kong, Zheng Du, Matthew Martin, and Tuan Vo-Dinh. Hyperspectral fluorescence image analysis for use in medical diagnostics. page 21, apr 2005.

[96] Seong G. Kong and Lae-Jeong Park. Hyperspectral Image Analysis for Skin Tumor Detection. In *Augmented Vision Perception in Infrared*, pages 155–171. Springer London, London, 2009.

[97] Farid Melgani and Lorenzo Bruzzone. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8):1778–1790, 2004.

[98] Zhi Liu, Jing-qi Yan, David Zhang, and Qing-Li Li. Automated tongue segmentation in hyperspectral images for medicine. *Applied optics*, 46(34):8328–8334, 2007.

[99] Khalid Masood and Nasir Rajpoot. Texture based classification of hyperspectral colon biopsy samples using clbp. In *Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009*, pages 1011–1014, 2009.

[100] Mathieu Fauvel and Jon Atli Benediktsson. Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profile. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11):3804–3814, 2008.

[101] Gordon F. Hughes. On the Mean Accuracy of Statistical Pattern Recognizers. *IEEE Transactions on Information Theory*, 14(1):55–63, 1968.

[102] Guolan Lu, Dongsheng Wang, Xulei Qin, Susan Muller, James V Little, Xu Wang, Amy Y Chen, Georgia Chen, and Baowei Fei. Histopathology feature mining and association with hyperspectral imaging for the detection of squamous neoplasia. *Scientific reports*, 9(1):1–13, 2019.

[103] George L Kumar and Lars Rudbeck. Immunohistochemical Staining Methods. Technical report, 2009.

[104] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621, 1973.

[105] Kevin J Chalut, Julie Hanson Ostrander, Michael G Giacomelli, and Adam Wax. Light scattering measurements of subcellular structure provide noninvasive early detection of chemotherapy-induced apoptosis. *Cancer research*, 69(3):1199–1204, 2009.

[106] Hariharan Subramanian, Prabhakar Pradhan, Yang Liu, Ilker R Capoglu, Xu Li, Jeremy D Rogers, Alexander Heifetz, Dhananjay Kunte, Hemant K Roy, Allen Taflove, et al. Optical methodology for detecting histologically unapparent nanoscale consequences of genetic alterations in biological cells. *Proceedings of the National Academy of Sciences*, 105(51):20118–20123, 2008.

[107] Rebekah A Drezek, Martial Guillaud, Thomas G Collier, Iouri Boiko, Anais Malpica, Calum E MacAulay, Michele Follen, and Rebecca R Richards-Kortum. Light scattering from cervical cells throughout neoplastic progression: influence of nuclear morphology, dna content, and chromatin texture. *Journal of biomedical optics*, 8(1):7–16, 2003.

[108] Anwer M. Siddiqi, Hui Li, Fazlay Faruque, Worth Williams, Kent Lai, Michael Hughson, Steven Bigler, James Beach, and William Johnson. Use of hyperspectral imaging to distinguish normal, precancerous, and cancerous cells. *Cancer*, 114(1):13–21, jan 2008.

[109] Niels Kröger-Lui, Norbert Gretz, Katharina Haase, Bettina Kränzlin, Sabine Neudecker, Annemarie Pucci, Alexander Regenscheit, Arthur Schönhals, and Wolfgang Petrich. Rapid identification of goblet cells in unstained colon thin sections by means of quantum cascade laser-based infrared microspectroscopy. *The Analyst*, 140(7):2086–92, apr 2015.

[110] Niels Kröger, Alexander Egl, Maria Engel, Norbert Gretz, Katharina Haase, Iris Herpich, Bettina Kränzlin, Sabine Neudecker, Annemarie Pucci, Arthur Schönhals, Jochen Vogt, and Wolfgang Petrich. Quantum cascade laser-based hyperspectral imaging of biological tissue. *Journal of Biomedical Optics*, 19(11):111607, jun 2014.

[111] Mang Ou-Yang, Yao Fang Hsieh, and Cheng Chung Lee. Biopsy diagnosis of oral carcinoma by the combination of morphological and spectral methods based on embedded relay lens microscopic hyperspectral imaging system. *Journal of Medical and Biological Engineering*, 35(4):437–447, 2015.

[112] Tuan Vo-Dinh. A hyperspectral imaging system for in vivo optical diagnostics. *IEEE Engineering in Medicine and Biology Magazine*, 23(5):40–49, 2004.

[113] Accelerated Access Review: Final Report. Review of innovative medicines and medical technologies. Accessed on 07-08-21.

[114] Martin Halicek, James D Dormer, James V Little, Amy Y Chen, Larry Myers, Baran D Sumer, and Baowei Fei. Hyperspectral imaging of head and neck squamous cell carcinoma for cancer margin detection in surgical specimens from 102 patients using deep learning. *Cancers*, 11(9):1367, 2019.

[115] Francesca Manni, Fons Van Der Sommen, Sveta Zinger, Esther Kho, Susan Brouwer De Koning, Theo Ruers, Caifeng Shan, Jean Schleipen, et al. Automated tumor assessment of squamous cell carcinoma on tongue cancer patients with hyperspectral imaging. In *Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 10951, page 109512K. International Society for Optics and Photonics, 2019.

[116] Zachary J Smith and Andrew J Berger. Integrated raman-and angular-scattering microscopy. *Optics letters*, 33(7):714–716, 2008.

[117] Zachary J Smith and Andrew J Berger. Validation of an integrated raman-and angular-scattering microscopy system on heterogeneous bead mixtures and single human immune cells. *Applied Optics*, 48(10):D109–D120, 2009.

[118] Jonathan Coppeta and Chris Rogers. Dual Emission Laser Induced Fluorescence For Direct Planar Scalar Behavior Measurements.

[119] Scott Prahl. Optical Absorption of Hemoglobin, 1999.

[120] Arunthathi Manickavasagam. *Multimodal Imaging of Optical Biomarkers for Label-free Oral Cancer Diagnosis.* King's College London, 2016.

[121] Susana Romero-Garcia, Jose Sullivan Lopez-Gonzalez, José Luis B´ ez Viveros, Dolores Aguilar-Cazares, and Heriberto Prado-Garcia. Tumor cell metabolism: an integral view. *Cancer biology & therapy*, 12(11):939–948, 2011.

[122] Manjul Tiwari. Science behind human saliva. *Journal of natural science, biology, and medicine*, 2(1):53, 2011.

[123] Iyad S Saidi, Steven L Jacques, and Frank K Tittel. Mie and rayleigh modeling of visible-light scattering in neonatal skin. *Applied optics*, 34(31):7410–7418, 1995.

[124] Maria Angela Franceschini, Enrico Gratton, Dennis Hueber, and Sergio Fantini. Near-Infrared Absorption and Scattering Spectra of Tissues in Vivo.

[125] Xiaobo Luo, Hao Xu, Mingjing He, Qi Han, Hui Wang, Chongkui Sun, Jing Li, Lu Jiang, Yu Zhou, Hongxia Dan, et al. Accuracy of autofluorescence in diagnosing oral squamous cell carcinoma and oral potentially malignant disorders: a comparative study with aero-digestive lesions. *Scientific reports*, 6(1):1–13, 2016.

[126] Elizabeth A Carter, Koman K Tam, Robert S Armstrong, and Peter A Lay. Vibrational spectroscopic mapping and imaging of tissues and cells. *Biophysical reviews*, 1(2):95–103, 2009.

[127] Stephan Preibisch, Stephan Saalfeld, and Pavel Tomancak. Globally optimal stitching of tiled 3d microscopic image acquisitions. *Bioinformatics*, 25(11):1463–1465, 2009.

[128] Imran I Patel, Júlio Trevisan, Geraint Evans, Valon Llabjani, Pierre L Martin-Hirsch, Helen F Stringfellow, and Francis L Martin. High contrast images of uterine tissue derived using raman microspectroscopy with the empty modelling approach of multivariate curve resolution-alternating least squares. *Analyst*, 136(23):4950–4959, 2011.

[129] Marcela Garcia Martinez, Anthony J Bullock, Sheila MacNeil, and Ihtesham U Rehman. Characterisation of structural changes in collagen with raman spectroscopy. *Applied Spectroscopy Reviews*, 54(6):509–542, 2019.

[130] Jesse Jaso and Reenu Malhotra. Adenoid cystic carcinoma. *Archives of pathology & laboratory medicine*, 135(4):511–515, 2011.

[131] Reinhard E Friedrich, Lei Li, Joerg Knop, Manfred Giese, and Rainer Schmelzle. Pleomorphic adenoma of the salivary glands: analysis of 94 patients. *Anticancer research*, 25(3A):1703–1705, 2005.

[132] Judith R Mourant, Andreas H Hielscher, Angelia A Eick, Tamara M Johnson, and James P Freyer. Evidence of intrinsic differences in the light scattering properties of tumorigenic and nontumorigenic cells. *Cancer Cytopathology: Interdisciplinary International Journal of the American Cancer Society*, 84(6):366–374, 1998.

[133] Madhuri Deolekar and J A Morris. How accurate are subjective judgements of a continuum? *Histopathology*, 42(3):227–232, 2003.

[134] Paola Taroni, Andrea Bassi, Daniela Comelli, Andrea Farina, Rinaldo Cubeddu, and Antonio Pifferi. Diffuse optical spectroscopy of breast tissue extended to 1100 nm. *Journal of biomedical optics*, 14(5):054030, 2009.

[135] Chrysovalantis Voutouri, Nathaniel D Kirkpatrick, Euiheon Chung, Fotios Mpekris, James W Baish, Lance L Munn, Dai Fukumura, Triantafyllos Stylianopoulos, and Rakesh K Jain. Experimental and computational analyses reveal dynamics of tumor vessel cooption and optimal treatment strategies. *Proceedings of the National Academy of Sciences*, 116(7):2662–2671, 2019.

[136] Martin Halicek, Himar Fabelo, Samuel Ortega, Gustavo M Callico, and Baowei Fei. In-vivo and ex-vivo tissue analysis through hyperspectral imaging techniques: revealing the invisible features of cancer. *Cancers*, 11(6):756, 2019.

[137] Johnny Gasperi, Stephanie L Wright, Rachid Dris, France Collard, Corinne Mandin, Mohamed Guerrouache, Valérie Langlois, Frank J Kelly, and Bruno Tassin. Microplastics in air: are we breathing it in? *Current Opinion in Environmental Science & Health*, 1:1–5, 2018.

[138] Ken Donaldson and C Lang Tran. Inflammation caused by particles and fibers. *Inhalation toxicology*, 14(1):5–27, 2002.

[139] Jie-Guang Chen and Susan Band Horwitz. Differential mitotic responses to microtubule-stabilizing and-destabilizing drugs. *Cancer research*, 62(7):1935–1938, 2002.

[140] Mary Ann Jordan, Kim Wendell, Sara Gardiner, W Brent Derry, Hillary Copp, and Leslie Wilson. Mitotic block induced in hela cells by low concentrations of paclitaxel (taxol) results in abnormal mitotic exit and apoptotic cell death. *Cancer research*, 56(4):816–825, 1996.

[141] Prasad V Jallepalli and Christoph Lengauer. Chromosome segregation and cancer: cutting through the mystery. *Nature Reviews Cancer*, 1(2):109–117, 2001.

[142] Judith R Mourant, Murat Canpolat, C Brocker, O Esponda-Ramos, Tamara M Johnson, A Matanock, K Stetter, and James P Freyer. Light scattering from cells: the contribution of the nucleus and the effects of proliferative status. *Journal of biomedical optics*, 5(2):131–137, 2000.

[143] Rekha Gautam, Sandeep Vanga, Freek Ariese, and Siva Umapathy. Review of multidimensional data processing approaches for raman and infrared spectroscopy. *EPJ Techniques and Instrumentation*, 2(1):1–38, 2015.

[144] Site Information for London Marylebone Road(UKA00315). https://uk-air.defra.gov.uk/networks . Accessed on 29-05-21.

[145] Camby Mei King Se, Kiao Inthavong, and Jiyuan Tu. Inhalability of micron particles through the nose and mouth. *Inhalation toxicology*, 22(4):287–300, 2010.

[146] Michael B Wallace, Lev T Perelman, Vadim Backman, James M Crawford, Marianne Fitzmaurice, Maurice Seiler, Kamran Badizadegan, Steven J Shields, Irving Itzkan, Ramachandra R Dasari, et al. Endoscopic detection of dysplasia in patients with barrett's esophagus using light-scattering spectroscopy. *Gastroenterology*, 119(3):677–682, 2000.

[147] Doddabasavaiah Basavapur Nandini and Ramadugula Venkata Subramanyam. Nuclear features in oral squamous cell carcinoma: A computer-assisted microscopic study. *Journal of oral and maxillofacial pathology: JOMFP*, 15(2):177, 2011.

[148] Karimollah Hajian-Tilaki. Receiver operating characteristic (roc) curve analysis for medical diagnostic test evaluation. *Caspian journal of internal medicine*, 4(2):627, 2013.

[149] Kjetil Søreide. Receiver-operating characteristic curve analysis in diagnostic, prognostic and predictive biomarker research. *Journal of clinical pathology*, 62(1):1–5, 2009.

[150] Alberto M Marchevsky, Ann E Walts, Birgit I Lissenberg-Witte, and Erik Thunnissen. Pathologists should probably forget about kappa. percent agreement, diagnostic specificity and related metrics provide more clinically applicable measures of interobserver variability. *Annals of Diagnostic Pathology*, 47:151561, 2020.

[151] Mary L McHugh. Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282, 2012.

[152] Froukje LJ Cals, Senada Koljenović, José A Hardillo, Robert J Baatenburg de Jong, Tom C Bakker Schut, and Gerwin J Puppels. Development and validation of raman spectroscopic classification models to discriminate tongue squamous cell carcinoma from non-tumorous tissue. *Oral oncology*, 60:41–47, 2016.

[153] Ioan Notingher. Raman spectroscopy cell-based biosensors. *sensors*, 7(8):1343–1358, 2007.

[154] Wei-Yun Dai, Szetsen Lee, and Yih-Chih Hsu. Discrimination between oral cancer and healthy cells based on the adenine signature detected by using raman spectroscopy. *Journal of Raman Spectroscopy*, 49(2):336–342, 2018.

[155] Baowei Fei, Guolan Lu, Xu Wang, Hongzheng Zhang, James V Little, Mihir R Patel, Christopher C Griffith, Mark W El-Diery, and Amy Y Chen. Label-free reflectance hyperspectral imaging for tumor margin assessment: a pilot study on surgical specimens of cancer patients. *Journal of biomedical optics*, 22(8):086009, 2017.

[156] Claus Kuepper, Angela Kallenbach-Thieltges, Hendrik Juette, Andrea Tannapfel, Frederik Großerueschkamp, and Klaus Gerwert. Quantum cascade laser-based infrared microscopy for label-free and automated cancer classification in tissue sections. *Scientific reports*, 8(1):1–10, 2018.

[157] Jean Guillaume Coutard, Mickael Brun, Maryse Fournier, Olivier Lartigue, Florian Fedeli, G Maisons, Jean-Marc Fedeli, Sergio Nicoletti, Mathieu Carras, and Laurent Duraffourg. Volume fabrication of quantum cascade lasers on 200 mm-cmos pilot line. *Scientific reports*, 10(1):1–8, 2020.

[158] Irving J Bigio and Stephen G Bown. Spectrroscopic sensing of cancer and cancer therapy: Current status of translational research. *Cancer biology & therapy*, 3(3):259–267, 2004.

[159] Stephen P Naber. Molecular pathology–detection of neoplasia. *New England Journal of Medicine*, 331(22):1508–1510, 1994.

[160] R Malini, Krishna Venkatakrishna, Jacob Kurien, Keerthilatha Pai, Lakshmi Rao, Vasudevan Bhaskaran Kartha, and Murali Krishna Chilakapati. Discrimination of normal, inflammatory, premalignant, and malignant oral tissue: a raman spectroscopy study. *Biopolymers: Original Research on Biomolecules*, 81(3):179–193, 2006.

[161] Yujuan Sun, Nan Liu, Xiaobing Guan, Hongru Wu, Zheng Sun, and Hui Zeng. Immunosuppression induced by chronic inflammation and the progression to oral squamous cell carcinoma. *Mediators of inflammation*, 2016, 2016.

[162] Zhiwei Huang, Annette McWilliams, Harvey Lui, David I McLean, Stephen Lam, and Haishan Zeng. Near-infrared raman spectroscopy for optical diagnosis of lung cancer. *International journal of cancer*, 107(6):1047–1052, 2003.

[163] Guolan Lu, James V Little, Xu Wang, Hongzheng Zhang, Mihir R Patel, Christopher C Griffith, Mark W El-Deiry, Amy Y Chen, and Baowei Fei. Detection of head and neck cancer in surgical specimens using quantitative hyperspectral imaging. *Clinical Cancer Research*, 23(18):5426–5436, 2017.

[164] Sanathana Konugolu Venkata Sekar, Ilaria Bargigia, Alberto Dalla Mora, Paola Taroni, Alessandro Ruggeri, Alberto Tosi, Antonio Pifferi, and Andrea Farina. Diffuse optical characterization of collagen absorption from 500 to 1700 nm. *Journal of biomedical optics*, 22(1):015006, 2017.

[165] Gabriela Botelho Martins, Sílvia Regina Reis, and Tânia Maria Silva. Collagen type i expression in squamous cell carcinoma of the oral cavity. *Pesquisa odontologica brasileira= Brazilian oral research*, 17(1):82–88, 2003.

[166] Priyanka Kardam, Monica Mehendiratta, Shweta Rehani, Madhumani Kumra, Khushboo Sahay, and Kanu Jain. Stromal fibers in oral squamous cell carcinoma: A possible new prognostic indicator? *Journal of oral and maxillofacial pathology: JOMFP*, 20(3):405, 2016.

[167] Rifat Hasina and Mark W Lingen. Angiogenesis in oral cancer. *Journal of dental education*, 65(11):1282–1290, 2001.

[168] Liqi Cai, Jundong Wang, Jinping Peng, Zhi Tan, Zhiwei Zhan, Xiangling Tan, and Qiuqiang Chen. Characteristic of microplastics in the atmospheric fallout from dongguan city, china: preliminary research and first evidence. *Environmental Science and Pollution Research*, 24(32):24928–24935, 2017.

[169] Rachid Dris, Johnny Gasperi, Mohamed Saad, Cécile Mirande, and Bruno Tassin. Synthetic fibers in atmospheric fallout: a source of microplastics in the environment? *Marine pollution bulletin*, 104(1-2):290–293, 2016.

[170] Stephanie L Wright, Joseph M Levermore, and Frank J Kelly. Raman spectral imaging for the detection of inhalable microplastics in ambient particulate matter samples. *Environmental science & technology*, 53(15):8947–8956, 2019.

[171] Laura Frère, Ika Paul-Pont, Julien Moreau, Philippe Soudant, Christophe Lambert, Arnaud Huvet, and Emmanuel Rinnert. A semi-automated raman micro-spectroscopy method for morphological and chemical characterizations of microplastic litter. *Marine pollution bulletin*, 113(1-2):461–468, 2016.

[172] Ali AlSuwaidi, Bruce Grieve, and Hujun Yin. Combining spectral and texture features in hyperspectral image analysis for plant monitoring. *Measurement Science and Technology*, 29(10):104001, 2018.

A Publications

Jarman, Anneliese, et al. "Hyperspectral microscopy and cluster analysis for oral cancer diagnosis." High-Speed Biomedical Imaging and Spectroscopy: Toward Big Data Instrumentation and Management II. Vol. 10076. International Society for Optics and Photonics, 2017.

Cited By

[1] Ortega, Samuel, et al. "Hyperspectral and multispectral imaging in digital and computational pathology: a systematic review." Biomedical Optics Express 11.6 (2020): 3195-3233.

[2] Ma, Ling, et al. "Hyperspectral microscopic imaging for automatic detection of head and neck squamous cell carcinoma using histologic image and machine learning." Medical Imaging 2020: Digital Pathology. Vol. 11320. International Society for Optics and Photonics, 2020.

[3] Wang, Jiansheng, and Qingli Li. "Quantitative analysis of liver tumors at different stages using microscopic hyperspectral imaging technology." Journal of biomedical optics 23.10 (2018): 106002.

B Open Source Software

A selection of the software solutions designed and employed in this project will be made available on the following webpage: https://github.com/a-jarman